GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 29, 2005, 06:23:18 ; Search time 11588 Seconds
(without alignments)
11436.401 Million cell updates/sec

Title: US-10-624-201A-1
Perfect score: 2735
Sequence: 1 catgcagagataaaaatata.....gcaaaaaaaaaaaaaaaaa 2735

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2735	100.0	2735	8	AF406697	AF406697 Solanum t
2	2135.8	78.1	2687	8	BT013459	BT013459 Lycopersi
3	2097.8	76.7	2636	8	AF375966	AF375966 Lycopersi
4	524.6	19.2	1898	8	AF406698	AF406698 Solanum t
5	401.2	14.7	2128	8	AF406702	AF406702 Solanum t
6	324.6	11.9	2043	6	AX506642	AX506642 Sequence
7	324.6	11.9	2043	8	BT008422	BT008422 Arabidops
8	324.6	11.9	2074	8	AY096548	AY096548 Arabidops
9	324.6	11.9	2385	8	AF353094	AF353094 Arabidops
10	324.6	11.9	2572	8	AY072175	AY072175 Arabidops
11	324.6	11.9	2578	8	AY099854	AY099854 Arabidops
12	295.6	10.8	1896	8	AK104872	AK104872 Oryza sat
13	295.6	10.8	2476	8	AK121356	AK121356 Oryza sat
14	295.6	10.8	3030	8	AK070465	AK070465 Oryza sat
15	295.2	10.8	2065	8	AF406703	AF406703 Solanum t
16	278.8	10.2	2523	8	AY524972	AY524972 Oryza sat
17	278.8	10.2	2570	8	AK067939	AK067939 Oryza sat
18	278.8	10.2	2630	8	AK069994	AK069994 Oryza sat
19	266.4	9.7	2952	8	AK070543	AK070543 Oryza sat
20	259.4	9.5	1449	6	AX412914	AX412914 Sequence
21	259.4	9.5	1449	8	BT000609	BT000609 Arabidops
22	259.4	9.5	2126	8	AY056796	AY056796 Arabidops
23	250.6	9.2	1731	8	AF406700	AF406700 Solanum t
24	249	9.1	1907	8	AF375964	AF375964 Lycopersi
25	241.8	8.8	1939	8	AF406701	AF406701 Solanum t
26	238	8.7	1840	8	AF406699	AF406699 Solanum t
27	234.4	8.6	1634	8	BT000133	BT000133 Arabidops
28	234.4	8.6	1824	8	AY085278	AY085278 Arabidops
29	234.4	8.6	1946	8	AY598452	AY598452 Arabidops
30	234.4	8.6	1950	8	AF353093	AF353093 Arabidops
31	234.4	8.6	1973	8	AY062545	AY062545 Arabidops
32	233.4	8.5	3097	8	AF053769	AF053769 Malus x d
33	230.2	8.4	1906	8	AY570508	AY570508 Arabidops
34	229.8	8.4	3132	8	GGN318871	AJ318871 Gnetum gn
35	225.4	8.2	1599	6	AX507459	AX507459 Sequence
36	225.4	8.2	1599	8	BT012291	BT012291 Arabidops
37	224	8.2	1884	6	AX506210	AX506210 Sequence
38	224	8.2	1884	8	AY142019	AY142019 Arabidops
39	224	8.2	2050	8	AF353092	AF353092 Arabidops
40	224	8.2	2124	8	AY058179	AY058179 Arabidops
41	222.6	8.1	2545	8	AF173816	AF173816 Arabidops
42	222.6	8.1	3162	8	AY050459	AY050459 Arabidops
43	219.4	8.0	28712	8	AC145458	AC145458 Cicer ari
44	204.8	7.5	2636	8	AK068034	AK068034 Oryza sat
45	204.8	7.5	2698	8	AK099829	AK099829 Oryza sat

ALIGNMENTS

RESULT 1
AF406697
LOCUS AF406697 2735 bp mRNA linear PLN 14-JUL-2003
DEFINITION Solanum tuberosum BEL1-related homeotic protein 5 (Bel5) mRNA, complete cds.
ACCESSION AF406697
VERSION AF406697.1 GI:22652114
KEYWORDS .
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiids; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 2735)
AUTHORS Chen,H., Rosin,F.M., Prat,S. and Hannapel,D.J.
TITLE Interacting transcription factors from the three-amino acid loop extension superclass regulate tuber formation
JOURNAL Plant Physiol. 132 (3), 1391-1404 (2003)
PUBMED 12857821
REFERENCE 2 (bases 1 to 2735)
AUTHORS Chen,H., Rosin,F.M. and Hannapel,D.J.
TITLE A KNOX protein of potato interacts with several members of the TALE family of transcription factors
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 2735)
AUTHORS Chen,H., Rosin,F.M. and Hannapel,D.J.
TITLE Direct Submission
JOURNAL Submitted (07-AUG-2001) Department of Horticulture, Iowa State University, 257 Horticulture Hall, Ames, IA 50011-1100, USA
FEATURES Location/Qualifiers
source 1. .2735
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Desiree"
/db_xref="taxon:4113"
gene 1. .2735
/gene="Bel5"
CDS 148. .2214
/gene="Bel5"
/note="StBel5"
/codon_start=1
/product="BEL1-related homeotic protein 5"

/protein_id="AAN03621.1"
/db_xref="GI:22652115"
/translation="MYYQGTSDNTNIQADHQQRHNHGNSNNNNIQTLYLMNPNNYMQG
YTTSDTQQQQQLLFLNSSPAASNALCHANIQHAPLQQQHFVGVPLPAVSLHDQINHHG
LLQRMWNNQDQSQQVIVPSSTGVSATSCGGITTDLASQLAFQRPIPTPQHRQQQQQQG
GLSLSLSPQLQQQISFNNNISSSSPRTNNVTIRGTLDGSSSNMVLGSKYLKAAQELLD
EVVNIVGKSIKGDDQKKDNSMNKESMPLASDVNTNSSGGGESSSRQKNEVAVELTTAQ
RQELQMKKAKLLAMLEEVEQRYRQYHHQMQIIVLSFEQVAGIGSAKSYTQLALHAISK
QFRCLKDAIAEQVKATSKSLGEEEGLGGKIEGSRLKFVDHHLRQQRALQQIGMMQPNA
WRPQRGLPERAVSVLRAWLFEHFLHPYPKDSDKIMLAKQTGLTRSQVSNWFINARVRL
WKPMVEEMYLEEVKNQEQNSTNTSGDNKNKETNISAPNEEKHPIITSSLLQDGITTTQ
AEISTSTISTSPTAGASLHHAHNFSFLGSFNMDNTTTTVDHIENNAKKQRNDMHKFSP
SSILSSVDMEAKARESSNKGFTNPLMAAYAMGDFGRFDPHDQQMTANFHGNNGVSLTL
GLPPSENLAMPVSQQNYLSNDLGSRSEMGSHYNRMGYENIDFQSGNKRFPTQLLPDFV
TGNLGT"

ORIGIN

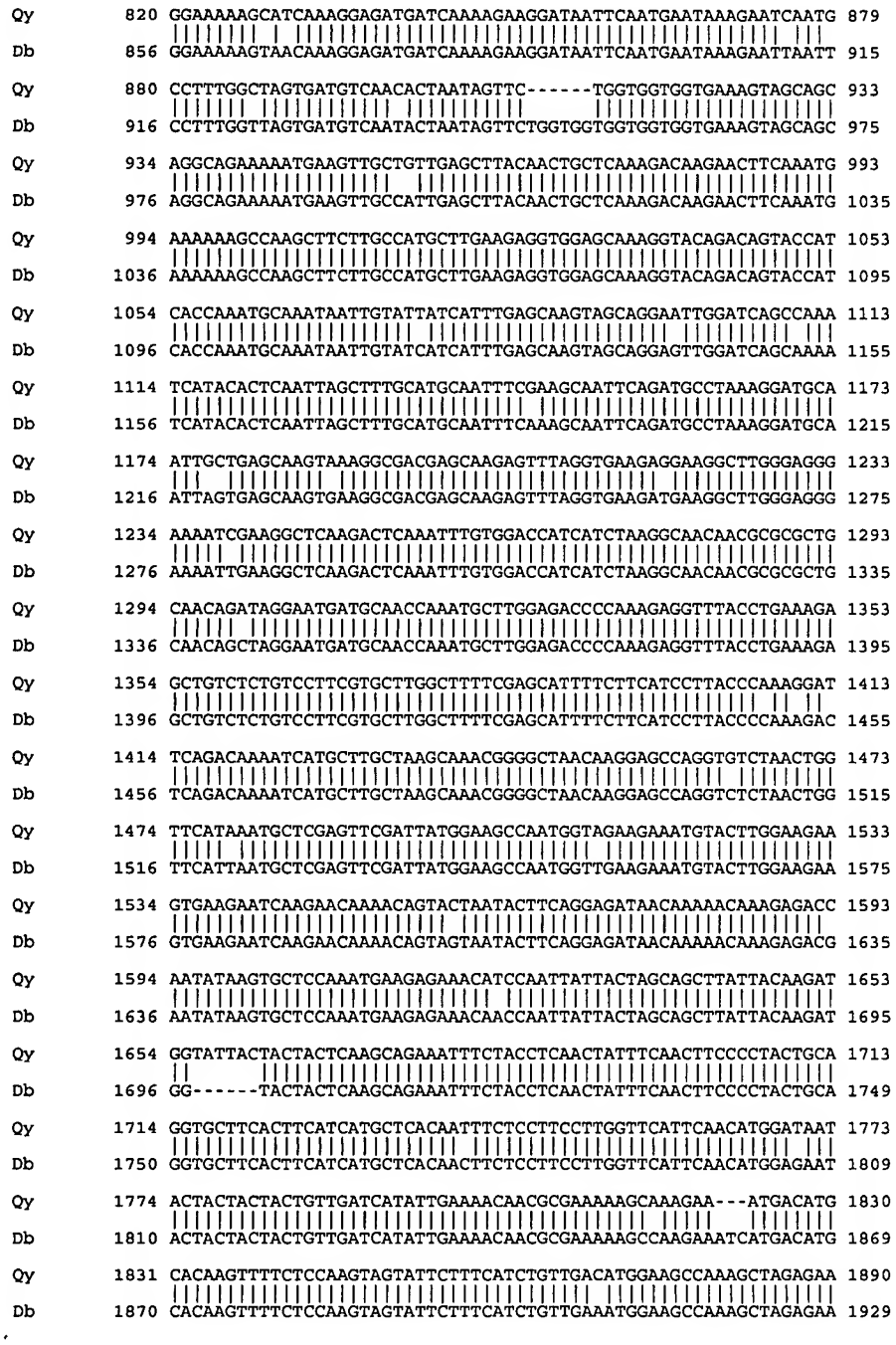
Query Match 100.0%; Score 2735; DB 8; Length 2735;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

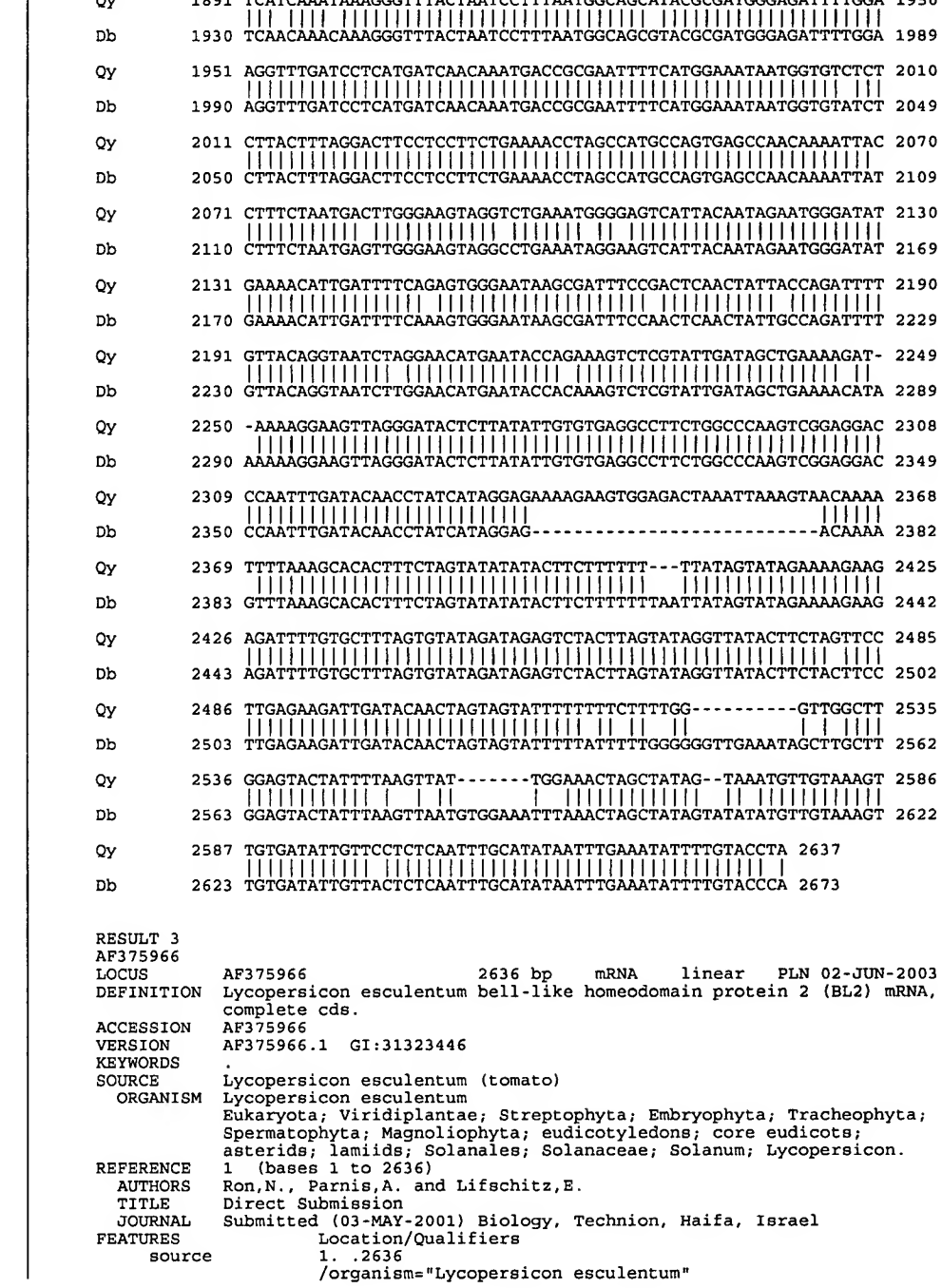
```
Qy          1 CATGCAGAGATAAAAATATAGATCAGTCTGACAAGAAGGCAACTTCTCAAAGCTTAGAGA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CATGCAGAGATAAAAATATAGATCAGTCTGACAAGAAGGCAACTTCTCAAAGCTTAGAGA 60

Qy         61 GCTACCACCCGAAGATAGACAGTTAGTTACATGTACTGTTATAGATAAAAGGAGAAATCC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCTACCACCCGAAGATAGACAGTTAGTTACATGTACTGTTATAGATAAAAGGAGAAATCC 120

Qy        121 GAAGAAGAAAGAATTTTTTTTTGCAGATATGTACTATCAAGGAACCTCGGATAATACTAAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAAGAAGAAAGAATTTTTTTTTGCAGATATGTACTATCAAGGAACCTCGGATAATACTAAT 180

Qy        181 ATACAAGCTGATCATCAACAACGTCATAATCATGGGAATAGTAATAATAATAATATTCAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATACAAGCTGATCATCAACAACGTCATAATCATGGGAATAGTAATAATAATAATATTCAG 240

Qy        241 ACACTTTATTTGATGAACCCTAACAATTATATGCAAGGCTACACTACTTCTGACACACAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACACTTTATTTGATGAACCCTAACAATTATATGCAAGGCTACACTACTTCTGACACACAG 300

Qy        301 CAGCAGCAGCAGTTACTTTTCCTGAATTCTTCACCAGCAGCAAGCAACGCGCTTTGCCAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGCAGCAGCAGTTACTTTTCCTGAATTCTTCACCAGCAGCAAGCAACGCGCTTTGCCAT 360

Qy        361 GCGAATATACAACACGCGCCGCTGCAACAGCAGCACTTTGTCGGTGTGCCTCTTCCGGCA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCGAATATACAACACGCGCCGCTGCAACAGCAGCACTTTGTCGGTGTGCCTCTTCCGGCA 420

Qy        421 GTAAGTTTGCACGATCAGATCAATCATCATGGACTTTTACAGCGCATGTGGAACAACCAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTAAGTTTGCACGATCAGATCAATCATCATGGACTTTTACAGCGCATGTGGAACAACCAA 480

Qy        481 GATCAATCTCAGCAGGTGATAGTACCATCGTCGACGGGGGTTTCTGCCACGTCATGTGGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GATCAATCTCAGCAGGTGATAGTACCATCGTCGACGGGGGTTTCTGCCACGTCATGTGGC 540

Qy        541 GGGATCACCACGGACTTGGCGTCTCAATTGGCGTTTCAGAGGCCGATTCCGACACCACAA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGGATCACCACGGACTTGGCGTCTCAATTGGCGTTTCAGAGGCCGATTCCGACACCACAA 600

Qy        601 CACCGACAGCAGCAACAACAGCAAGGCGGTCTATCTCTAAGCCTTTCTCCTCAGCTACAA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CACCGACAGCAGCAACAACAGCAAGGCGGTCTATCTCTAAGCCTTTCTCCTCAGCTACAA 660

Qy        661 CAGCAAATTAGTTTCAATAACAATATTTCATCCTCATCACCAAGGACAAATAATGTTACT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CAGCAAATTAGTTTCAATAACAATATTTCATCCTCATCACCAAGGACAAATAATGTTACT 720

Qy        721 ATTAGGGGAACATTAGATGGAAGTTCTAGCAACATGGTTTTAGGCTCTAAGTATCTGAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATTAGGGGAACATTAGATGGAAGTTCTAGCAACATGGTTTTAGGCTCTAAGTATCTGAAA 780

Qy        781 GCTGCACAAGAGCTTCTTGATGAAGTTGTTAATATTGTTGGAAAAAGCATCAAAGGAGAT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GCTGCACAAGAGCTTCTTGATGAAGTTGTTAATATTGTTGGAAAAAGCATCAAAGGAGAT 840

Qy        841 GATCAAAAGAAGGATAATTCAATGAATAAAGAATCAATGCCTTTGGCTAGTGATGTCAAC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GATCAAAAGAAGGATAATTCAATGAATAAAGAATCAATGCCTTTGGCTAGTGATGTCAAC 900

Qy        901 ACTAATAGTTCTGGTGGTGGTGAAAGTAGCAGCAGGCAGAAAAATGAAGTTGCTGTTGAG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 ACTAATAGTTCTGGTGGTGGTGAAAGTAGCAGCAGGCAGAAAAATGAAGTTGCTGTTGAG 960

Qy        961 CTTACAACTGCTCAAAGACAAGAACTTCAAATGAAAAAAGCCAAGCTTCTTGCCATGCTT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CTTACAACTGCTCAAAGACAAGAACTTCAAATGAAAAAAGCCAAGCTTCTTGCCATGCTT 1020

Qy       1021 GAAGAGGTGGAGCAAAGGTACAGACAGTACCATCACCAAATGCAAATAATTGTATTATCA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GAAGAGGTGGAGCAAAGGTACAGACAGTACCATCACCAAATGCAAATAATTGTATTATCA 1080

Qy       1081 TTTGAGCAAGTAGCAGGAATTGGATCAGCCAAATCATACACTCAATTAGCTTTGCATGCA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TTTGAGCAAGTAGCAGGAATTGGATCAGCCAAATCATACACTCAATTAGCTTTGCATGCA 1140

Qy       1141 ATTTCGAAGCAATTCAGATGCCTAAAGGATGCAATTGCTGAGCAAGTAAAGGCGACGAGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ATTTCGAAGCAATTCAGATGCCTAAAGGATGCAATTGCTGAGCAAGTAAAGGCGACGAGC 1200

Qy       1201 AAGAGTTTAGGTGAAGAGGAAGGCTTGGGAGGGAAAATCGAAGGCTCAAGACTCAAATTT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AAGAGTTTAGGTGAAGAGGAAGGCTTGGGAGGGAAAATCGAAGGCTCAAGACTCAAATTT 1260

Qy       1261 GTGGACCATCATCTAAGGCAACAACGCGCGCTGCAACAGATAGGAATGATGCAACCAAAT 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GTGGACCATCATCTAAGGCAACAACGCGCGCTGCAACAGATAGGAATGATGCAACCAAAT 1320

Qy       1321 GCTTGGAGACCCCAAAGAGGTTTACCTGAAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GCTTGGAGACCCCAAAGAGGTTTACCTGAAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTT 1380

Qy       1381 TTCGAGCATTTTCTTCATCCTTACCCAAAGGATTCAGACAAAATCATGCTTGCTAAGCAA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TTCGAGCATTTTCTTCATCCTTACCCAAAGGATTCAGACAAAATCATGCTTGCTAAGCAA 1440

Qy       1441 ACGGGGCTAACAAGGAGCCAGGTGTCTAACTGGTTCATAAATGCTCGAGTTCGATTATGG 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 ACGGGGCTAACAAGGAGCCAGGTGTCTAACTGGTTCATAAATGCTCGAGTTCGATTATGG 1500

Qy       1501 AAGCCAATGGTAGAAGAAATGTACTTGGAAGAAGTGAAGAATCAAGAACAAAACAGTACT 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 AAGCCAATGGTAGAAGAAATGTACTTGGAAGAAGTGAAGAATCAAGAACAAAACAGTACT 1560

Qy       1561 AATACTTCAGGAGATAACAAAAACAAAGAGACCAATATAAGTGCTCCAAATGAAGAGAAA 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AATACTTCAGGAGATAACAAAAACAAAGAGACCAATATAAGTGCTCCAAATGAAGAGAAA 1620

Qy       1621 CATCCAATTATTACTAGCAGCTTATTACAAGATGGTATTACTACTACTCAAGCAGAAATT 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 CATCCAATTATTACTAGCAGCTTATTACAAGATGGTATTACTACTACTCAAGCAGAAATT 1680

Qy       1681 TCTACCTCAACTATTTCAACTTCCCCTACTGCAGGTGCTTCACTTCATCATGCTCACAAT 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 TCTACCTCAACTATTTCAACTTCCCCTACTGCAGGTGCTTCACTTCATCATGCTCACAAT 1740

Qy       1741 TTCTCCTTCCTTGGTTCATTCAACATGGATAATACTACTACTACTGTTGATCATATTGAA 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TTCTCCTTCCTTGGTTCATTCAACATGGATAATACTACTACTACTGTTGATCATATTGAA 1800

Qy       1801 AACAACGCGAAAAAGCAAAGAAATGACATGCACAAGTTTTCTCCAAGTAGTATTCTTTCA 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 AACAACGCGAAAAAGCAAAGAAATGACATGCACAAGTTTTCTCCAAGTAGTATTCTTTCA 1860

Qy       1861 TCTGTTGACATGGAAGCCAAAGCTAGAGAATCATCAAATAAAGGGTTTACTAATCCTTTA 1920
```

```
Db       1861 TCTGTTGACATGGAAGCCAAAGCTAGAGAATCATCAAATAAAGGGTTTACTAATCCTTTA 1920

Qy       1921 ATGGCAGCATACGCGATGGGAGATTTTGGAAGGTTTGATCCTCATGATCAACAAATGACC 1980
Db       1921 ATGGCAGCATACGCGATGGGAGATTTTGGAAGGTTTGATCCTCATGATCAACAAATGACC 1980

Qy       1981 GCGAATTTTCATGGAAATAATGGTGTCTCTCTTACTTTAGGACTTCCTCCTTCTGAAAAC 2040
Db       1981 GCGAATTTTCATGGAAATAATGGTGTCTCTCTTACTTTAGGACTTCCTCCTTCTGAAAAC 2040

Qy       2041 CTAGCCATGCCAGTGAGCCAACAAAATTACCTTTCTAATGACTTGGGAAGTAGGTCTGAA 2100
Db       2041 CTAGCCATGCCAGTGAGCCAACAAAATTACCTTTCTAATGACTTGGGAAGTAGGTCTGAA 2100

Qy       2101 ATGGGGAGTCATTACAATAGAATGGGATATGAAAACATTGATTTTCAGAGTGGGAATAAG 2160
Db       2101 ATGGGGAGTCATTACAATAGAATGGGATATGAAAACATTGATTTTCAGAGTGGGAATAAG 2160

Qy       2161 CGATTTCCGACTCAACTATTACCAGATTTTGTTACAGGTAATCTAGGAACATGAATACCA 2220
Db       2161 CGATTTCCGACTCAACTATTACCAGATTTTGTTACAGGTAATCTAGGAACATGAATACCA 2220

Qy       2221 GAAAGTCTCGTATTGATAGCTGAAAAGATAAAAGGAAGTTAGGGATACTCTTATATTGTG 2280
Db       2221 GAAAGTCTCGTATTGATAGCTGAAAAGATAAAAGGAAGTTAGGGATACTCTTATATTGTG 2280

Qy       2281 TGAGGCCTTCTGGCCCAAGTCGGAGGACCCAATTTGATACAACCTATCATAGGAGAAAAG 2340
Db       2281 TGAGGCCTTCTGGCCCAAGTCGGAGGACCCAATTTGATACAACCTATCATAGGAGAAAAG 2340

Qy       2341 AAGTGGAGACTAAATTAAAGTAACAAAATTTTAAAGCACACTTTCTAGTATATATACTTC 2400
Db       2341 AAGTGGAGACTAAATTAAAGTAACAAAATTTTAAAGCACACTTTCTAGTATATATACTTC 2400

Qy       2401 TTTTTTTTATAGTATAGAAAAGAAGAGATTTTGTGCTTTAGTGTATAGATAGAGTCTACT 2460
Db       2401 TTTTTTTTATAGTATAGAAAAGAAGAGATTTTGTGCTTTAGTGTATAGATAGAGTCTACT 2460

Qy       2461 TAGTATAGGTTATACTTCTAGTTCCTTGAGAAGATTGATACAACTAGTAGTATTTTTTTT 2520
Db       2461 TAGTATAGGTTATACTTCTAGTTCCTTGAGAAGATTGATACAACTAGTAGTATTTTTTTT 2520

Qy       2521 CTTTTGGGTTGGCTTGGAGTACTATTTTAAGTTATTGGAAACTAGCTATAGTAAATGTTG 2580
Db       2521 CTTTTGGGTTGGCTTGGAGTACTATTTTAAGTTATTGGAAACTAGCTATAGTAAATGTTG 2580

Qy       2581 TAAAGTTGTGATATTGTTCCTCTCAATTTGCATATAATTTGAAATATTTTGTACCTACTA 2640
Db       2581 TAAAGTTGTGATATTGTTCCTCTCAATTTGCATATAATTTGAAATATTTTGTACCTACTA 2640

Qy       2641 GCTAGTCTCTAAATTATGTTTCCATTGCTTGTAATTGCAATTTTATTTGAATTTTGTGCT 2700
Db       2641 GCTAGTCTCTAAATTATGTTTCCATTGCTTGTAATTGCAATTTTATTTGAATTTTGTGCT 2700

Qy       2701 ATCATTATTAGATTAGCAAAAAAAAAAAAAAAAAA 2735
Db       2701 ATCATTATTAGATTAGCAAAAAAAAAAAAAAAAAA 2735
```

```
RESULT 2
BT013459
LOCUS       BT013459               2687 bp    mRNA    linear   PLN 11-MAY-2004
DEFINITION  Lycopersicon esculentum clone 132117F, mRNA sequence.
ACCESSION   BT013459
VERSION     BT013459.1  GI:47104874
KEYWORDS    FLI_CDNA.
SOURCE      Lycopersicon esculentum (tomato)
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1  (bases 1 to 2687)
  AUTHORS   Kirkness,E.F., Wang,W. and Vazeille,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-MAY-2004) The Institute for Genomic Research, 9712
            Medical Center Drive, Rockville, MD 20850, USA
FEATURES             Location/Qualifiers
     source          1. .2687
                     /organism="Lycopersicon esculentum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4081"
                     /clone="132117F"
                     /tissue_type="fruit"
                     /note="TZXBC46"
ORIGIN
```

```
  Query Match             78.1%;  Score 2135.8;  DB 8;  Length 2687;
  Best Local Similarity   91.1%;  Pred. No. 0;
  Matches 2452;  Conservative    0;  Mismatches  132;  Indels  107;  Gaps   13;

Qy         12 AAAAATATAGATCAGTCTGACAAGAAGGCAACTTCTCAAAGCTTAGAGAGCTACCACCCG 71
Db         25 AAAAATATAGATCAGTCTGACAAGAAGGCAACTTCTCAAAGCTTAGAGAGCTACCACCCA 84

Qy         72 AAGATAGACAGTTAGTTACA--TGTACTGTTATAGATAAAAGGAGAAATCCGAAGAAGAA 129
Db         85 AAGATAGACAGTTAGTTACATGTGTACTGTTATTGATAAAAGGAGAAATTTGAACAAG-- 142

Qy        130 AGAATTTTTTTTGCAGATATGTACTATCAAGGAACCTCGGATAATACTAATATACAAGCT 189
Db        143 -------TTTGTGGAGATATGTACTATCAAGGAACCTCGGATAATAATATACAAGCTGAT 195

Qy        190 GATCATCAACAACGTCATAATCATGGGAATAGTAATAATAATAATATTCAGACACTTTAT 249
Db        196 CATCATCAACAACAACATAATAATCTTGGGAATAGTAATAATAATATTCAGACGCTGTAT 255

Qy        250 TTGATGAACCCTAACAATTATATGCAAGGCTACACTACTTCTGACACACAGCAG------ 303
Db        256 CTGATGAACCCTAACAGTTATATGCAAGGCTACACTACAACTGACACACAGCAGCACCTG 315

Qy        304 --------------CAGCAGCAGTTACTTTTCCTGAATTCTTCACCAGCAGCAAGCAAC 348
Db        316 CAACAGCAACAGAACCAGCACCAGTTGCTTTTTCTGAATTCAGCACCAGCAGGAGGCAAC 375

Qy        349 GCGCTTTGCCATGCGAATATACAACACGCGCCGCTGCAACAGCAGCACTTTGTCGGTGTG 408
Db        376 GCGCTTAGCCATGCGAATATACAACACGCGCCTCTGCAACAGCAGCACTTTGTCGGCGTG 435

Qy        409 CCTCTTCCGGCAGTAAGTTTGCACGATCAGATCAATCATCATGGACTTTTACAGCGCATG 468
Db        436 CCCCTTCCGGCAGTGAGTTTGCATGATCAGATCAATCATCACGGACTTTTACAGCGTATG 495

Qy        469 TGGAACAACCAAGATCAATCTCAGCAGGTGATAGTACCATCGTCGACGGGGGTTTCTGCC 528
Db        496 TGGAACAATCAAGACCAATCTCAGCAGGTGATAGTACCGTCGTCGACGGTGGTTTCTGCC 555

Qy        529 ACGTCATGTGGCGGGATCACCACGGACTTGGCGTCTCAATTGGCGTTTCAGAGGCCGAT- 587
Db        556 ACGTCATGTGGCGGGACCACCACGGACTTGGCGTCTCAACTGGCGTTTCAGAGGCCGATT 615

Qy        588 --------TCCGACACCACAACACCGACAGCAGCAACAACAGCAAGGCGGTCTATCTCTA 639
Db        616 GTGGTGTCTCCGACACCACAACACCGACAACAACAACAACAGCAAGGCGGTCTATCACTG 675

Qy        640 AGCCTTTCTCCTCAGCTACAACAGCAAATTAGTTTCAATAACAATATTTCATCCTCATCA 699
Db        676 AGCCTTTCTCCTCAGCAACAGCAACAAATTAGTTTCAATAACAATATTTCATCATCATCA 735

Qy        700 CCAAGGACAAATAATGTTACTATTAGAGGGAACATTAGATGGAAGTTCTAGCAACATGGTT 759
Db        736 CCAAGGACAAATAATGTTACTATTAGAGGAACAATGGATGGATGTTCTAGCAACATGATT 795

Qy        760 TTAGGCTCTAAGTATCTGAAAGCTGCACAAGAGCTTCTTGATGAAGTTGTTAATATTGTT 819
Db        796 TTAGGCTCTAAGTATCTGAAAGCTGCACAAGAGCTTCTTGATGAAGTGGTTAATATTGTT 855
```





```
RESULT 3
AF375966
LOCUS       AF375966               2636 bp    mRNA    linear   PLN 02-JUN-2003
DEFINITION  Lycopersicon esculentum bell-like homeodomain protein 2 (BL2) mRNA,
            complete cds.
ACCESSION   AF375966
VERSION     AF375966.1  GI:31323446
KEYWORDS    .
SOURCE      Lycopersicon esculentum (tomato)
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1  (bases 1 to 2636)
  AUTHORS   Ron,N., Parnis,A. and Lifschitz,E.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-MAY-2001) Biology, Technion, Haifa, Israel
FEATURES             Location/Qualifiers
     source          1. .2636
                     /organism="Lycopersicon esculentum"
```

```
                     /mol_type="mRNA"
                     /db_xref="taxon:4081"
     gene            1. .2636
                     /gene="BL2"
     CDS             100. .2199
                     /gene="BL2"
                     /note="TBL2"
                     /codon_start=1
                     /product="bell-like homeodomain protein 2"
                     /protein_id="AAP47025.1"
                     /db_xref="GI:31323447"
                     /translation="MYYQGTSDNNIQADHHQQQHNNLGNSNNNIQTLYLMNPNSYMQG
                     YTTTDTQQHLQQQQNQHQLLFLNSAPAGGNALSHANIQHAPLQQQHFVGVPLPAVSLH
                     DQINHHGLLQRMWNNQDQSQQVIVPSSTVVSATSCGGTTTDLASQLAFQRPIVVSPTP
                     QHRQQQQQQGGLSLSLSPQQQQQISFNNNISSSSPRTNNVTIRGTMDGCSSNMILGSK
                     YLKAAQELLDEVVNIVGKSNKGDDQKKDNSMNKELIPLVSDVNTNSSGGGGGESSSRQ
                     KNEVAIELTTAQRQELQMKKAKLLAMLEEVEQRYRQYHHQMQIIVSSFEQVAGVGSAK
                     SYTQLALHAISKQFRCLKDAISEQVKATSKSLGEDEGLGGKIEGSRLKFVDHHLRQQR
                     ALQQLGMMQPNAWRPQRGLPERAVSVLRAWLFEHFLHPYPKDSDKIMLAKQTGLTRSQ
                     VSNWFINARVRLWKPMVEEMYLEEVKNQEQNSSNTSGDNKNETNISAPNEEKQPIIT
                     SSLLQDGTTQAEISTSTISTSPTAGASLHHAHNFSFLGSFNMENTTTTVDHIENNAKK
                     PRNHDMHKFSPSSILSSVEMEAKARESTNKGFTNPLMAAYAMGDFGRFDPHDQQMTAN
                     FHGNNGVSLTLGLPPSENLAMPVSQQNYLSNELGSRPEIGSHYNRMGYENIDFQSGNK
                     RFPTQLLPDFVTGNLGT"
ORIGIN

  Query Match           76.7%;  Score 2097.8;  DB 8;  Length 2636;
  Best Local Similarity 90.9%;  Pred. No. 0;
  Matches 2418;  Conservative    0;  Mismatches  132;  Indels  111;  Gaps   13;

Qy         46 CTCAAAGCTTAGAGAGCTACCACCCGAAGATAGACAGTTAGTTACAT------GTACTGT 99
              ||||||||||||||||||||||||| ||||||||||||||||||||      |||||||
Db          1 CTCAAAGCTTAGAGAGCTACCACCCAAAGATAGACAGTTAGTTACATGTGTGAGTACTGT 60

Qy        100 TATAGATAAAAGGAGAAATCCGAAGAAGAAAGAATTTTTTTTGCAGATATGTACTATCAA 159
              ||| ||||||||||||||  ||| |||          ||| || ||||||||||||||||
Db         61 TATTGATAAAAGGAGAAATTTGAACAAG---------TTTGTGGAGATATGTACTATCAA 111

Qy        160 GGAACCTCGGATAATACTAATATACAAGCTGATCATCAACAACGTCATAATCATGGGAAT 219
              |||||||||||||||| ||     |  | ||||||||||||  ||||| ||
Db        112 GGAACCTCGGATAATAATATACAAGCTGATCATCATCAACAACAACATAATAATCTTGGG 171

Qy        220 AGTAATAATAATAATATTCAGACACTTTATTTGATGAACCCTAACAATTATATGCAAGGC 279
              | || |||||||||||||||||| || ||| |||||||||||||| ||||||||||||||
Db        172 AATAGTAATAATAATATTCAGACGCTGTATCTGATGAACCCTAACAGTTATATGCAAGGC 231

Qy        280 TACACTACTTCTGACACACAGCAG----------------------CAGCAGCAGTTACTT 318
              ||||||||  ||||||||||||||                     ||||| ||||| |||
Db        232 TACACTACAACTGACACACAGCAGCACCTGCAACAGCAACAGAACCAGCACCAGTTGCTT 291

Qy        319 TTCCTGAATTCTTCACCAGCAGCAAGCAACGCGCTTTGCCATGCGAATATACAACACGCG 378
              || ||||||||  |||||||||  | ||||||||||  ||||||||||||||||||||||
Db        292 TTTCTGAATTCAGCACCAGCAGGAGGCAACGCGCTTAGCCATGCGAATATACAACACGCG 351

Qy        379 CCGCTGCAACAGCAGCACTTTGTCGGTGTGCCTCTTCCGGCAGTAAGTTTGCACGATCAG 438
              || |||||||||||||||||||||| ||||| |||||||||| |||||||| ||||||
Db        352 CCTCTGCAACAGCAGCACTTTGTCGGCGTGCCCCTTCCGGCAGTGAGTTTGCATGATCAG 411

Qy        439 ATCAATCATCATGGACTTTTACAGCGCATGTGGAACAACCAAGATCAATCTCAGCAGGTG 498
              |||||||||| |||||||||||||| |||||||||| ||||| ||||||||||||||||
Db        412 ATCAATCATCACGGACTTTTACAGCGTATGTGGAACAATCAAGACCAATCTCAGCAGGTG 471

Qy        499 ATAGTACCATCGTCGACGGGGGTTTCTGCCACGTCATGTGGCGGGATCACCACGGACTTG 558
              |||||||| |||||||||| ||||||||||||||||||||||||| ||||||||||||||
Db        472 ATAGTACCGTCGTCGACGGTGGTTTCTGCCACGTCATGTGGCGGGACCACCACGGACTTG 531

Qy        559 GCGTCTCAATTGGCGTTTCAGAGGCCGAT---------TCCGACACCACAACACCGACAG 609
              ||||||||| ||||||||||||||||||         |||||||||||||||||||||
Db        532 GCGTCTCAACTGGCGTTTCAGAGGCCGATTGTGGTGTCTCCGACACCACAACACCGACAA 591

Qy        610 CAGCAACAACAGCAAGGCGGTCTATCTCTAAGCCTTTCTCCTCAGCTACAACAGCAAATT 669
              || |||||||||||||| |||||||| || |||||||||||||||| ||| || ||||||
Db        592 CAACAACAACAGCAAGGGGGTCTATCACTGAGCCTTTCTCCTCAGCAACAGCAACAAATT 651

Qy        670 AGTTTCAATAACAATATTTCATCCTCATCACCAAGGACAAATAATGTTACTATTAGGGGA 729
              ||||||||||||||||||||||| ||||||||||||||||||||||||||||||| |||
Db        652 AGTTTCAATAACAATATTTCATCATCATCACCAAGGACAAATAATGTTACTATTAGAGGA 711

Qy        730 ACATTAGATGGAAGTTCTAGCAACATGGTTTTAGGCTCTAAGTATCTGAAAGCTGCACAA 789
              ||| | |||||| |||||||||||||| ||||||||||||||||||||||||||||||
Db        712 ACAATGGATGGATGTTCTAGCAACATGATTTTAGGCTCTAAGTATCTGAAAGCTGCACAA 771

Qy        790 GAGCTTCTTGATGAAGTTGTTAATATTGTTGGAAAAAGCATCAAAGGAGATGATCAAAAG 849
              |||||||||||||||| ||||||||||||||||||| | ||||||||||||||||||
Db        772 GAGCTTCTTGATGAAGTGGTTAATATTGTTGGAAAAAGTAACAAAGGAGATGATCAAAAG 831

Qy        850 AAGGATAATTCAATGAATAAAGAATCAATGCCTTTGGCTAGTGATGTCAACACTAATAGT 909
              |||||||||||||||||||||||| ||| ||||||| |||||||||||| |||||||||
Db        832 AAGGATAATTCAATGAATAAAGAATTAATTCCTTTGGTTAGTGATGTCAATACTAATAGT 891

Qy        910 TC------TGGTGGTGGTGAAAGTAGCAGCAGGCAGAAAAATGAAGTTGCTGTTGAGCTT 963
              ||      |||||||||||||||||||||||||||||||||||||||||| ||||||||
Db        892 TCTGGTGGTGGTGGTGGTGAAAGTAGCAGCAGGCAGAAAAATGAAGTTGCCATTGAGCTT 951

Qy        964 ACAACTGCTCAAAGACAAGAACTTCAAATGAAAAAAGCCAAGCTTCTTGCCATGCTTGAA 1023
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        952 ACAACTGCTCAAAGACAAGAACTTCAAATGAAAAAAGCCAAGCTTCTTGCCATGCTTGAA 1011

Qy       1024 GAGGTGGAGCAAAGGTACAGACAGTACCATCACCAAATGCAAATAATTGTATTATCATTT 1083
              ||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db       1012 GAGGTGGAGCAAAGGTACAGACAGTACCATCACCAAATGCAAATAATTGTATCATCATTT 1071

Qy       1084 GAGCAAGTAGCAGGAATTGGATCAGCCAAATCATACACTCAATTAGCTTTGCATGCAATT 1143
              ||||||||||||||| |||||||||| |||||||||||||||||||||||||||||||||
Db       1072 GAGCAAGTAGCAGGAGTTGGATCAGCAAAATCATACACTCAATTAGCTTTGCATGCAATT 1131

Qy       1144 TCGAAGCAATTCAGATGCCTAAAGGATGCAATTGCTGAGCAAGTAAAGGCGACGAGCAAG 1203
              || |||||||||||||||||||||||||||||| ||||||||| ||||||||||||||||
Db       1132 TCAAAGCAATTCAGATGCCTAAAGGATGCAATTAGTGAGCAAGTGAAGGCGACGAGCAAG 1191

Qy       1204 AGTTTAGGTGAAGAGGAAGGCTTGGGAGGGAAAAATCGAAGGCTCAAGACTCAAATTTGTG 1263
              |||||||||||||| |||||||||||||||||||| ||||||||||||||||||||||||
Db       1192 AGTTTAGGTGAAGATGAAGGCTTGGGAGGGAAAATTGAAGGCTCAAGACTCAAATTTGTG 1251

Qy       1264 GACCATCATCTAAGGCAACAACGCGCGCTGCAACAGATAGGAATGATGCAACCAAATGCT 1323
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1252 GACCATCATCTAAGGCAACAACGCGCGCTGCAACAGCTAGGAATGATGCAACCAAATGCT 1311

Qy       1324 TGGAGACCCCAAAGAGGTTTACCTGAAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTTTTC 1383
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1312 TGGAGACCCCAAAGAGGTTTACCTGAAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTTTTC 1371

Qy       1384 GAGCATTTTCTTCATCCTTACCCAAAGGATTCAGACAAAATCATGCTTGCTAAGCAAACG 1443
              |||||||||||||||||||||| || || ||||||||||||||||||||||||||||||
Db       1372 GAGCATTTTCTTCATCCTTACCCCAAAGACTCAGACAAAATCATGCTTGCTAAGCAAACG 1431

Qy       1444 GGGCTAACAAGGAGCCAGGTGTCTAACTGGTTCATAAATGCTCGAGTTCGATTATGGAAG 1503
              |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db       1432 GGGCTAACAAGGAGCCAGGTCTCTAACTGGTTCATTAATGCTCGAGTTCGATTATGGAAG 1491

Qy       1504 CCAATGGTAGAAGAAATGTACTTGGAAGAAGTGAAGAATCAAGAACAAAACAGTACTAAT 1563
              |||||||| ||||||||||||||||||||||||||||||||||||||||||| ||||
Db       1492 CCAATGGTTGAAGAAATGTACTTGGAAGAAGTGAAGAATCAAGAACAAAACAGTAGTAAT 1551

Qy       1564 ACTTCAGGAGATAACAAAAACAAAGAGACCAATATAAGTGCTCCAAATGAAGAGAAACAT 1623
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db       1552 ACTTCAGGAGATAACAAAAACAAAGAGACGAATATAAGTGCTCCAAATGAAGAGAAACAA 1611

Qy       1624 CCAATTATTACTAGCAGCTTATTACAAGATGGTATTACTACTACTCAAGCAGAAATTTCT 1683
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1612 CCAATTATTACTAGCAGCTTATTACAAGATGG------TACTACTCAAGCAGAAATTTCT 1665

Qy       1684 ACCTCAACTATTTCAACTTCCCCTACTGCAGGTGCTTCACTTCATCATGCTCACAATTTC 1743
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db       1666 ACCTCAACTATTTCAACTTCCCCTACTGCAGGTGCTTCACTTCATCATGCTCACAACTTC 1725
```

```
Qy       1744 TCCTTCCTTGGTTCATTCAACATGGATAATACTACTACTACTGTTGATCATATTGAAAAC 1803
             ||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db       1726 TCCTTCCTTGGTTCATTCAACATGGAGAATACTACTACTACTGTTGATCATATTGAAAAC 1785

Qy       1804 AACGCGAAAAAGCAAAGAA---ATGACATGCACAAGTTTTCTCCAAGTAGTATTCTTTCA 1860
             ||||||||||||| |||||   ||||||||||||||||||||||||||||||||||||||||
Db       1786 AACGCGAAAAAGCCAAGAAATCATGACATGCACAAGTTTTCTCCAAGTAGTATTCTTTCA 1845

Qy       1861 TCTGTTGACATGGAAGCCAAAGCTAGAGAATCATCAAATAAAGGGTTTACTAATCCTTTA 1920
             |||||||| ||||||||||||||||||||| |||| ||||||||||||||||||||||||
Db       1846 TCTGTTGAAATGGAAGCCAAAGCTAGAGAATCAACAAACAAAGGGTTTACTAATCCTTTA 1905

Qy       1921 ATGGCAGCATACGCGATGGGAGATTTTGGAAGGTTTGATCCTCATGATCAACAAATGACC 1980
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1906 ATGGCAGCGTACGCGATGGGAGATTTTGGAAGGTTTGATCCTCATGATCAACAAATGACC 1965

Qy       1981 GCGAATTTTCATGGAAATAATGGTGTCTCTCTTACTTTAGGACTTCCTCCTTCTGAAAAC 2040
             |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db       1966 GCGAATTTTCATGGAAATAATGGTGTATCTCTTACTTTAGGACTTCCTCCTTCTGAAAAC 2025

Qy       2041 CTAGCCATGCCAGTGAGCCAACAAAATTACCTTTCTAATGACTTGGGAAGTAGGTCTGAA 2100
             |||||||||||||||||||||||||||| |||||||||||| |||||||||||| |||||
Db       2026 CTAGCCATGCCAGTGAGCCAACAAAATTATCTTTCTAATGAGTTGGGAAGTAGGCCTGAA 2085

Qy       2101 ATGGGGAGTCATTACAATAGAATGGGATATGAAAACATTGATTTTCAGAGTGGGAATAAG 2160
             || || ||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db       2086 ATAGGAAGTCATTACAATAGAATGGGATATGAAAACATTGATTTTCAAAGTGGGAATAAG 2145

Qy       2161 CGATTTCCGACTCAACTATTACCAGATTTTGTTACAGGTAATCTAGGAACATGAATACCA 2220
             |||||||| |||||||||| |||||||||||||||||||| ||||||||||||||||||||
Db       2146 CGATTTCCAACTCAACTATTGCCAGATTTTGTTACAGGTAATCTTGGAACATGAATACCA 2205

Qy       2221 GAAAGTCTCGTATTGATAGCTGAAAAGAT--AAAAGGAAGTTAGGGATACTCTTATATTG 2278
              |||||||||||||||||||||||| ||  ||||||||||||||||||||||||||||||
Db       2206 CAAAGTCTCGTATTGATAGCTGAAAACATAAAAAAGGAAGTTAGGGATACTCTTATATTG 2265

Qy       2279 TGTGAGGCCTTCTGGCCCAAGTCGGAGGACCCAATTTGATACAACCTATCATAGGAGAAA 2338
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2266 TGTGAGGCCTTCTGGCCCAAGTCGGAGGACCCAATTTGATACAACCTATCATAGGAG--- 2322

Qy       2339 AGAAGTGGAGACTAAATTAAAGTAACAAAATTTTAAAGCACACTTTCTAGTATATATACT 2398
                                      ||||||||||||||||||||||||||||||||||
Db       2323 ----------------------ACAAAAGTTTAAAGCACACTTTCTAGTATATATACT 2358

Qy       2399 TCTTTTTT---TTATAGTATAGAAAAGAAGAGATTTTGTGCTTTAGTGTATAGATAGAGT 2455
             ||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2359 TCTTTTTTTAATTATAGTATAGAAAAGAAGAGATTTTGTGCTTTAGTGTATAGATAGAGT 2418

Qy       2456 CTACTTAGTATAGGTTATACTTCTAGTTCCTTGAGAAGATTGATACAACTAGTAGTATTT 2515
             |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db       2419 CTACTTAGTATAGGTTATACTTCTACTTCCTTGAGAAGATTGATACAACTAGTAGTATTT 2478

Qy       2516 TTTTTCTTTTGG----------GTTGGCTTGGAGTACTATTTTAAGTTAT-------TGG 2558
             || || ||  ||          | | ||||||||||||||| |  | ||       |
Db       2479 TTATTTTTGGGGGGTTGAAATAGCTTGCTTGGAGTACTATTTAAGTTAATGTGGAAATTT 2538

Qy       2559 AAACTAGCTATAG--TAAATGTTGTAAAGTTGTGATATTGTTCCTCTCAATTTGCATATA 2616
             |||||||||||||  || |||||||||||||||||||| ||||||||||||||||||||||
Db       2539 AAACTAGCTATAGTATATATGTTGTAAAGTTGTGATATTGTTACTCTCAATTTGCATATA 2598

Qy       2617 ATTTGAAATATTTTGTACCTA 2637
             |||||||||||||||||||||
Db       2599 ATTTGAAATATTTTGTACCTA 2619


RESULT 4
AF406698
LOCUS       AF406698               1898 bp    mRNA    linear   PLN 14-JUL-2003
DEFINITION  Solanum tuberosum BEL1-related homeotic protein 11 (Bel11) mRNA,
            partial cds.
ACCESSION   AF406698
VERSION     AF406698.1  GI:22652116
KEYWORDS    .
SOURCE      Solanum tuberosum (potato)
  ORGANISM  Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum.
REFERENCE   1  (bases 1 to 1898)
  AUTHORS   Chen,H., Rosin,F.M., Prat,S. and Hannapel,D.J.
  TITLE     Interacting transcription factors from the three-amino acid loop
            extension superclass regulate tuber formation
  JOURNAL   Plant Physiol. 132 (3), 1391-1404 (2003)
   PUBMED   12857821
REFERENCE   2  (bases 1 to 1898)
  AUTHORS   Chen,H., Rosin,F.M. and Hannapel,D.J.
  TITLE     A KNOX protein of potato interacts with several members of the TALE
            family of transcription factors
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 1898)
  AUTHORS   Chen,H., Rosin,F.M. and Hannapel,D.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-AUG-2001) Department of Horticulture, Iowa State
            University, 257 Horticulture Hall, Ames, IA 50011-1100, USA
FEATURES             Location/Qualifiers
     source          1. .1898
                     /organism="Solanum tuberosum"
                     /mol_type="mRNA"
                     /cultivar="Desiree"
                     /db_xref="taxon:4113"
     gene            <1. .1898
                     /gene="Bel11"
     CDS             <1. .1608
                     /gene="Bel11"
                     /note="StBel11"
                     /codon_start=1
                     /product="BEL1-related homeotic protein 11"
                     /protein_id="AAN03622.1"
                     /db_xref="GI:22652117"
                     /translation="MTFRSSLPLDLREISTTNHQVGILSSSPLPSPGTNTNNINHTRG
                     LGASSSFSISNGMILGSKYLKVAQDLLDEVVNVGKNIKLSDGLESGAKEKHKLDNELI
                     SLASDDVESSSQKNSGVELTTAQRQELQMKKAKLVSMLDEVDQRYRQYHHQMQMIATS
                     FEQTTGIGSSKSYTQLALHTISKQFRCLKDAISGQIKDTSKTLGEEENIGGKIEGSKL
                     KFVDHHLRQQRALQQLGMMQTNAWKPQRGLPERAVSVLRAWLFEHFLHPYPKDSDKII
                     LAKQTGLTRSQVSNWFINARVRLWKPMVEEMYMEEVKKNNQEQNIEPNNNEIVGSKSS
                     VPQEKLPISSNIIHNASPNDISTSTISTSPTGGGGSIPTQTVAGFSFIRSLNMENIDD
                     QRNNKKARNEMQNCSTSTILSMEREIINKVVQDETIKSEKFNNTQTRECYSLMTPNYT
                     MDDQFGTRFNNQNHEQLATTTTFHQGNGHVSLTLGLPPNSENQHNYIGLENHYNQPTH
                     HPNISYENIDFQSGKRYATQLLQDFVS"
ORIGIN

  Query Match           19.2%;  Score 524.6;  DB 8;  Length 1898;
  Best Local Similarity  71.3%;  Pred. No. 7.5e-96;
  Matches  813;  Conservative    0;  Mismatches  274;  Indels   54;  Gaps    7;

Qy        729 AACATTAGATGGAAGTTCTAGCAACATGGTTTTAGGCTCTAAGTATCTGAAAGCTGCACA 788
              | |||    |   | |||||     ||| | || || |||||||| || |||| ||||||
Db        144 ATCATCTTTTTCGATTTCTAATGGGATGATATTGGGTTCTAAGTACCTAAAAGTTGCACA 203

Qy        789 AGAGCTTCTTGATGAAGTTGTTAATATTGTTGGAAAAAGCATCAAAGGAGATGATCAAAA 848
              ||| ||||||||||||||||||   |||||||||| |||||||  |   |||
Db        204 AGATCTTCTTGATGAAGTTGTTAA---TGTTGGAAAAAACATCAAATTATCAGATGGCTT 260

Qy        849 GAAGGATAATTCAATGAATAAAGAATCAATGCCTTTGGCTAGTGATGTCAACACTAATAG 908
              ||  |  | ||| | | ||| |    |      ||||  | |||  | |   ||
Db        261 AGAGAGTGGTGCAAAGGAGAAACACAAA------TTGGACAATGAATTAATATCTTTGGC 314

Qy        909 TTCTGGTGGTGGTGAAAGTAGCAGCAGGCAGAAAAATGAAGTTGCTGTTGAGCTTACAAC 968
              |  || || || |||||| ||||||       ||||  |   || |||||| ||||||||
Db        315 TAGTGATGATGTTGAAAGCAGCAGC------CAAAAAAATAGTGGTGTTGAACTTACAAC 368

Qy        969 TGCTCAAAGACAAGAACTTCAAATGAAAAAAGCCAAGCTTCTTGCCATGCTTGAAGAGGT 1028
              |||||||||||||||||||||||| |||||||||||| ||  ||||||||| |||||
Db        369 AGCTCAAAGACAAGAACTTCAAATGAAGAAAGCCAAGCTTGTTAGCATGCTTGATGAGGT 428

Qy       1029 GGAGCAAAGGTACAGACAGTACCATCACCAAATGCAAATAATTGTATTATCATTTGAGCA 1088
```

```
Db        429 GGATCAAAGGTATAGACAATACCATCACCAAATGCAAATGATTGCAACATCATTTGAGCA 488

Qy       1089 AGTAGCAGGAATTGGATCAGCCAAATCATACACTCAATTAGCTTTGCATGCAATTTCGAA 1148

Db        489 AACAACAGGAATTGGATCATCAAAATCATACACACAACTTGCTTTGCACACAATTTCAAA 548

Qy       1149 GCAATTCAGATGCCTAAAGGATGCAATTGCTGAGCAAGTAAAGGCGACGAGCAAGAGTTT 1208

Db        549 GCAATTTAGATGTTTAAAAGATGCAATTTCTGGGCAAATAAAGGACACTAGCAAAACTTT 608

Qy       1209 AGGTGAAGAGGAAGGCTTGGGAGGGAAAATCGAAGGCTCAAGACTCAAATTTGTGGACCA 1268

Db        609 AGGGGAAGAAGAAAACATTGGAGGCAAAATTGAAGGATCAAAGTTGAAATTTGTGGATCA 668

Qy       1269 TCATCTAAGGCAACAACGCGCGCTGCAACAGATAGGAATGATGCAACCAAATGCTTGGAG 1328

Db        669 TCATTTACGCCAACAACGTGCACTACAACAATTAGGGATGATGCAAACCAATGCATGGAA 728

Qy       1329 ACCCCAAAGAGGTTTACCTGAAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTTTTCGAGCA 1388

Db        729 GCCTCAAAGAGGTTTGCCAGAAAGAGCGGTTTCAGTTCTCCGCGCTTGGCTTTTCGAGCA 788

Qy       1389 TTTTCTTCATCCTTACCCAAAGGATTCAGACAAAATCATGCTTGCTAAGCAAACGGGGCT 1448

Db        789 TTTTCTTCATCCGTATCCCAAAGATTCAGATAAAATCATCCTTGCTAAGCAAACAGGGCT 848

Qy       1449 AACAAGGAGCCAGGTGTCTAACTGGTTCATAAATGCTCGAGTTCGATTATGGAAGCCAAT 1508

Db        849 AACAAGGAGCCAGGTATCAAATTGGTTTATAAATGCTAGAGTTAGACTATGGAAGCCAAT 908

Qy       1509 GGTAGAAGAAATGTACTTGGAAGAAGTGAAGAATCAAGAACAAAACAGTACTAATACTTC 1568

Db        909 GGTAGAAGAAATGTACATGGAAGAAGTGAAGAAAAACAATCAAGAACAAAATATTGAGCC 968

Qy       1569 AGGAGATAACAAAAACAAAGAGACCAATATAAGTGCTCCAAATGAAGAGAAACATCCAAT 1628

Db        969 TAATAACAATGAAATTGTTGGCTCAAAATCAAGTGTTCCA---CAAGAGAAATTACCAAT 1025

Qy       1629 TATTACTAGCAGCTTATTACAAGATGGTATTACTACTACTCAAGCAGAAATTTCTACCTC 1688

Db       1026 TAGTAGCA------------ATATTATTCATAATGCTTCTCCAAATGATATTTCTACTTC 1073

Qy       1689 AACTATTTCAACTTCCCCTACTGCAGGTG---------CTTCACTTCATCATGCTCACAA 1739

Db       1074 CACCATTTCAACATCTCCGACGGGTGGCGGCGGTTCGATTCCGACTCAGACGGTTGCAGG 1133

Qy       1740 TTTCTCCTTCCTTGGTTCATTCAACATGGATAATACTACTACTACTGTTGATCATATTGA 1799

Db       1134 TTTCTCCTTCATTAGGTCATTAAACATGGAGAACA--------------TTGATGATCA 1178

Qy       1800 AAACAACGCGAAAAAGCAAAGAAATGACATGCACAAGTTTTCTCCAAGTAGTATTCTTTC 1859

Db       1179 AAGGAACAACAAAAAGGCAAGAAATGAGATGCAAAATTGTTCAACTAGTACTATTCTCTC 1238

Qy       1860 A 1860

Db       1239 A 1239
```

```
RESULT 5
AF406702
LOCUS       AF406702               2128 bp    mRNA    linear   PLN 14-JUL-2003
DEFINITION  Solanum tuberosum BEL1-related homeotic protein 29 (Bel29) mRNA,
            partial cds.
ACCESSION   AF406702
VERSION     AF406702.1  GI:22652124
KEYWORDS    .
SOURCE      Solanum tuberosum (potato)
  ORGANISM  Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum.
REFERENCE   1  (bases 1 to 2128)
  AUTHORS   Chen,H., Rosin,F.M., Prat,S. and Hannapel,D.J.
  TITLE     Interacting transcription factors from the three-amino acid loop
            extension superclass regulate tuber formation
  JOURNAL   Plant Physiol. 132 (3), 1391-1404 (2003)
   PUBMED   12857821
REFERENCE   2  (bases 1 to 2128)
  AUTHORS   Chen,H., Rosin,F.M. and Hannapel,D.J.
  TITLE     A KNOX protein of potato interacts with several members of the TALE
            family of transcription factors
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 2128)
  AUTHORS   Chen,H., Rosin,F.M. and Hannapel,D.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-AUG-2001) Department of Horticulture, Iowa State
            University, 257 Horticulture Hall, Ames, IA 50011-1100, USA
FEATURES             Location/Qualifiers
     source          1. .2128
                     /organism="Solanum tuberosum"
                     /mol_type="mRNA"
                     /cultivar="Desiree"
                     /db_xref="taxon:4113"
     gene            <1. .2128
                     /gene="Bel29"
     CDS             <1. .1704
                     /gene="Bel29"
                     /note="StBel29"
                     /codon_start=1
                     /product="BEL1-related homeotic protein 29"
                     /protein_id="AAN03626.1"
                     /db_xref="GI:22652125"
                     /translation="QGLSLSLSSSQQPGFGNFTAARELVSSPSGSASASGIQQQQQQQ
                     QSISSVPLSSKYMKAAQELLDEVVNVGKSMKSTNSTDVVVNNDVKKSKNMGDMDGQLD
                     GVGADKDGAPTTELSTGERQEIQMKKAKLVNMLDEVEQRYRHYHHQMQSVIHWLEQAA
                     GIGSAKTYTALALQTISKQFRCLKDAIIGQIRSASQTLGEEDSLGGKIEGSRLKFVDN
                     QLRQQRALQQLGMIQHNAWRPQRGLPERAVSVLRAWLFEHFLHPYPKDSDKMMLAKQT
                     GLTRSQVSNWFINARVRLWKPMVEEMYLEEIKEHEQNGLGQEKTSKLGEQNEDSTTSR
                     SIATQDKSPGSDSQNKSFVSKQDNHLPQHNPASPMPDVQRHFHTPIGMTIRNQSAGFN
                     LIGSPEIESINITQGSPKKPRNNEMLHSPNSIPSINMDVKPNEEQMSMKFGDDRQDRD
                     GFSLMGGPMNFMGGFGAYPIGEIARFSTEQFSAPYSTSGTVSLTLGLPHNENLSMSAT
                     HHSFLPIPTQNIQIGSEPNHEFGSLNTPTSAHSTSSVYETFNIQNRKRFAAPLLPDFV
                     A"
ORIGIN
```

```
  Query Match             14.7%;  Score 401.2;  DB 8;  Length 2128;
  Best Local Similarity   70.2%;  Pred. No. 7.6e-71;
  Matches  570;  Conservative    0;  Mismatches  233;  Indels    9;  Gaps    2;

Qy        742 AGTTCTAGCAACATGGTTTTAGGCTCTAAGTATCTGAAAGCTGCACAAGAGCTTCTTGAT 801

Db        136 AGTATTAGTAGTGTGCCTTTGAGTTCTAAGTACATGAAGGCTGCACAAGAGCTACTTGAT 195

Qy        802 GAAGTTGTTAATATTGTTGGAAAAAGCATCAAAGGAGATGATCAAAAGAAGGATAATTCA 861

Db        196 GAAGTTGTAAA---TGTTGGAAAATCAATGAAAAGTACTAATAGTACTGATGTTGTTGTT 252

Qy        862 ATGAATAAAGAATCAATGCCTTTGGCTAGTGATGTCAACACTAATAGTTCTGGTGGTGGT 921

Db        253 AATAATGATGTCAAGAAATCGAAGAATATGGGCGATATGGACGGACAGTTAGACGGAGTT 312

Qy        922 GAAAGTAGCAGCAGGCAGAAAAATGAAGTTGCTGTTGAGCTTACAACTGCTCAAAGACAA 981

Db        313 GGAGCAGACA------AAGACGGAGCTCCAACAACTGAGCTAAGTACAGGGGAGAGACAA 366

Qy        982 GAACTTCAAATGAAAAAAGCCAAGCTTCTTGCCATGCTTGAAGAGGTGGAGCAAAGGTAC 1041

Db        367 GAAATTCAAATGAAGAAAGCAAAACTTGTTAACATGCTTGACGAGGTGGAGCAGAGGTAT 426

Qy       1042 AGACAGTACCATCACCAAATGCAAATAATTGTATTATCATTTGAGCAAGTAGCAGGAATT 1101

Db        427 AGACATTATCATCACCAAATGCAGTCAGTGATACATTGGTTAGAGCAAGCTGCTGGCATT 486

Qy       1102 GGATCAGCCAAATCATACACTCAATTAGCTTTGCATGCAATTTCGAAGCAATTCAGATGC 1161
```

```
Db        487 GGATCAGCAAAAACATATACAGCATTGGCTTTGCAGACGATTTCGAAGCAATTTAGGTGT 546

Qy       1162 CTAAAGGATGCAATTGCTGAGCAAGTAAAGGCGACGAGCAAGAGTTTAGGTGAAGAGGAA 1221

Db        547 CTTAAGGACGCGATAATTGGTCAAATACGATCAGCAAGCCAGACGTTAGGCGAAGAAGAT 606

Qy       1222 GGCTTGGGAGGGAAAATCGAAGGCTCAAGACTCAAATTTGTGGACCATCATCTAAGGCAA 1281

Db        607 AGTTTGGGAGGGAAGATTGAAGGTTCAAGGCTTAAATTTGTTGATAATCAGCTAAGACAG 666

Qy       1282 CAACGCGCGCTGCAACAGATAGGAATGATGCAACCAAATGCTTGGAGACCCCAAAGAGGT 1341

Db        667 CAAAGGGCTTTGCAACAATTGGGAATGATCCAGCATAATGCTTGGAGACCTCAGAGAGGA 726

Qy       1342 TTACCTGAAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTTTTCGAGCATTTTCTTCATCCT 1401

Db        727 TTGCCCGAACGAGCTGTTTCTGTTCTTCGCGCTTGGCTTTTTGAACATTTCCTCCATCCT 786

Qy       1402 TACCCAAAGGATTCAGACAAAATCATGCTTGCTAAGCAAACGGGGCTAACAAGGAGCCAG 1461

Db        787 TATCCCAAGGATTCAGACAAAATGATGCTAGCAAAACAAACAGGACTAACTAGGAGTCAG 846

Qy       1462 GTGTCTAACTGGTTCATAAATGCTCGAGTTCGATTATGGAAGCCAATGGTAGAAGAAATG 1521

Db        847 GTGTCGAATTGGTTCATCAATGCTCGAGTTCGTCTTTGGAAGCCAATGGTGGAAGAGATG 906

Qy       1522 TACTTGGAAGAAGTGAAGAATCAAGAACAAAA 1553

Db        907 TACTTGGAAGAGATAAAAGAACACGAACAGAA 938
```

RESULT 6
AX506642

LOCUS	AX506642 2043 bp DNA linear PAT 27-SEP-2002
DEFINITION	Sequence 1337 from Patent WO0216655.
ACCESSION	AX506642
VERSION	AX506642.1 GI:23387879
KEYWORDS	.
SOURCE	Arabidopsis thaliana (thale cress)
ORGANISM	Arabidopsis thaliana Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE	1
AUTHORS	Harper,J.F., Kreps,J., Wang,X. and Zhu,T.
TITLE	Stress-regulated genes of plants, transgenic plants containing same, and methods of use
JOURNAL	Patent: WO 0216655-A 1337 28-FEB-2002; The Scripps Research Institute (US) ; Syngenta Participations AG (CH)

FEATURES	Location/Qualifiers
source	1. .2043
	/organism="Arabidopsis thaliana"
	/mol_type="unassigned DNA"
	/db_xref="taxon:3702"

ORIGIN

Query Match 11.9%; Score 324.6; DB 6; Length 2043;
Best Local Similarity 65.2%; Pred. No. 2.5e-55;
Matches 534; Conservative 0; Mismatches 264; Indels 21; Gaps 3;

```
Qy        758 TTTTAGGCTCTAAGTATCTGAAAGCTGCACAAGAGCTTCTTGATGAAGTTGTTAATATTG 817

Db        569 TTGTTAGCTCCAAGTACTTGAAGGCAGCACAAGAGCTTCTTGACGAAGTAGTCAACGCTG 628

Qy        818 TTGGAAAAAGCATCAAAGGAGATGATCAAAAGAAGGATAATTCAATGAATAAAGAATCAA 877

Db        629 ATTCCGATGACATGAACGCTAAATCCCAACTATTCTCATCGAAAAAGGGTAGTTGCGGAA 688

Qy        878 TGCCTTTGGCTAGTGATGTCAACACTAATAGTTCTGGTGGTGGTGAAAGTAGCAGCAGGC 937

Db        689 ATGATAAACCTGTCGGAGAATCATCGGCCGGCGCTGGAGGAGAAGGTTCCGGTGGCGGAG 748
```

```
Qy        938 AGAAAAATG------AAGTTGCTGTTGAGCTTACAACTGCTCAAAGACAAGAACTTCAAA 991

Db        749 CAGAAGCAGCCGGGAAACGTCCGGTGGAGCTAGGCACGGCAGAGAGACAAGAAATACAGA 808

Qy        992 TGAAAAAAGCCAAGCTTCTTGCCATGCTTGAAGAGGTGGAGCAAAGGTACAGACAGTACC 1051

Db        809 TGAAGAAAGCAAAACTTAGTAACATGCTTCATGAGGTGGAGCAGAGATATAGACAGTACC 868

Qy       1052 ATCACCAAATGCAAATAATTGTATTATCATTTGAGCAAGTAGCAGGAATTGGATCAGCCA 1111

Db        869 ACCAGCAGATGCAGATGGTGATCTCTTCGTTCGAGCAAGCGGCAGGGATAGGATCAGCGA 928

Qy       1112 AATCATACACTCAATTAGCTTTGCATGCAATTTCGAAGCAATTCAGATGCCTAAAGGATG 1171

Db        929 AGTCATACACGTCGCTAGCATTGAAAACCATATCAAGACAGTTCCGTTGCTTGAAAGAGG 988

Qy       1172 CAATTGCTGAGCAAGTAAAGGCGACGAGCAAGAGTTTAGGTGAAGAGGAAGGCTTG---- 1227

Db        989 CGATCGCTGGTCAGATAAAAGCGGCCAACAAGAGTCTTGGGGAGGAAGATTCAGTGTCTG 1048

Qy       1228 --GGAGGGAAAATCGAAGGCTCAAGACTCAAATTTGTGGACCATCATCTAAGGCAACAAC 1285

Db        1049 GTGTTGGGAGGTTTGAGGGGTCGAGGCTCAAGTTCGTGGACCACCACTTGAGACAGCAAA 1108

Qy       1286 GCGCGCTGCAACAGATAGGAATGATGCAAC---------CAAATGCTTGGAGACCCCAAA 1336

Db        1109 GAGCTCTTCAACAACTGGGAATGATTCAACATCCTTCCAATAATGCTTGGAGACCTCAAC 1168

Qy       1337 GAGGTTTACCTGAAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTTTTCGAGCATTTTCTTC 1396

Db        1169 GTGGTCTCCCAGAACGAGCCGTCTCAGTTCTCCGTGCTTGGCTCTTCGAACACTTTCTTC 1228

Qy       1397 ATCCTTACCCAAAGGATTCAGACAAAATCATGCTTGCTAAGCAAACGGGGCTAACAAGGA 1456

Db        1229 ATCCATACCCTAAGGATTCGGACAAGCACATGCTAGCTAAGCAAACAGGACTCACTCGTA 1288

Qy       1457 GCCAGGTGTCTAACTGGTTCATAAATGCTCGAGTTCGATTATGGAAGCCAATGGTAGAAG 1516

Db        1289 GCCAGGTGTCGAACTGGTTTATAAACGCGAGAGTTCGGTTATGGAAACCAATGGTGGAGG 1348

Qy       1517 AAATGTACTTGGAAGAAGTGAAGAATCAAGAACAAAACA 1555

Db        1349 AGATGTACATGGAGGAAATGAAGGAGCAGGCAAAGAACA 1387
```

RESULT 7
BT008422

LOCUS	BT008422 2043 bp mRNA linear PLN 15-MAY-2003
DEFINITION	Arabidopsis thaliana At2g35940 gene, complete cds.
ACCESSION	BT008422
VERSION	BT008422.1 GI:30725517
KEYWORDS	FLI_CDNA.
SOURCE	Arabidopsis thaliana (thale cress)
ORGANISM	Arabidopsis thaliana Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE	1 (bases 1 to 2043)
AUTHORS	Cheuk,R., Chen,H., Kim,C.J., Shinn,P., Bowser,L., Carninci,P., Dale,J.M., Hayashizaki,Y., Hsuan,V.W., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Toriumi,M., Wong,C., Wu,H.C., Yamada,K., Yu,G., Yuan,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.
TITLE	Arabidopsis ORF clones
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 2043)
AUTHORS	Cheuk,R., Chen,H., Kim,C.J., Shinn,P., Bowser,L., Carninci,P., Dale,J.M., Hayashizaki,Y., Hsuan,V.W., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L.,

```
            Sakurai,T., Satou,M., Seki,M., Southwick,A., Toriumi,M., Wong,C.,
            Wu,H.C., Yamada,K., Yu,G., Yuan,S., Shinozaki,K., Davis,R.W.,
            Theologis,A. and Ecker,J.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-MAY-2003) Salk Institute Genomic Analysis Laboratory
            (SIGnAL), Plant Biology Laboratory, The Salk Institute for
            Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
            USA
COMMENT     RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
            Arabidopsis Full-Length cDNA') :  Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P.,  Kawai,J.,
            Hayashizaki,Y. and Shinozaki,K.

            The Salk, Stanford, PGEC (SSP) Consortium members constructed and
            sequenced the pUNI (ORF) clones using the RAFL cDNAs: Cheuk,R.,
            Chen,H., Kim,C.J., Shinn,P., Bowser,L., Chan, M.M., Chang,C.M.,
            Dale,J.M., Hsuan,V.W., Jones,T., Karlin-Neumann,G., Lam,B.,
            Lee,J.M., Lin,J., Miranda,M., Nguyen,M., Onodera, C.S., Palm,C.J.,
            Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C.,
            Yamada,K., Yu,G., Yuan,S., Davis,R.W.,  Theologis,A., and
            Ecker,J.R.

            Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to
            this work.  Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
            contributed equally to this work as PIs.
FEATURES              Location/Qualifiers
     source           1. .2043
                      /organism="Arabidopsis thaliana"
                      /mol_type="mRNA"
                      /db_xref="taxon:3702"
                      /chromosome="2"
                      /clone="U25364"
                      /ecotype="Columbia"
                      /note="This clone is in pUNI 51"
     CDS              1. .2043
                      /note="putative homeodomain transcription factor"
                      /codon_start=1
                      /product="At2g35940"
                      /protein_id="AAP37781.1"
                      /db_xref="GI:30725518"
                      /translation="MAAYFHGNPPEISAGSDGGLQTLILMNPTTYVQYTQQDNDSNNN
                      NNSNNSNNNNTNTNTNNNNSSFVFLDSHAPQPNASQQFVGIPLSGHEAASITAADNIS
                      VLHGYPPRVQYSLYGSHQVDPTHQQAACETPRAQQGLSLTLSSQQQQQQQHHQQHQPI
                      HVGFGSGHGEDIRVGSGSTGSGVTNGIANLVSSKYLKAAQELLDEVVNADSDDMNAKS
                      QLFSSKKGSCGNDKPVGESSAGAGGEGSGGGAEAAGKRPVELGTAERQEIQMKKAKLS
                      NMLHEVEQRYRQYHQQMQMVISSFEQAAGIGSAKSYTSLALKTISRQFRCLKEAIAGQ
                      IKAANKSLGEEDSVSGVGRFEGSRLKFVDHHLRQQRALQQLGMIQHPSNNAWRPQRGL
                      PERAVSVLRAWLFEHFLHPYPKDSDKHMLAKQTGLTRSQVSNWFINARVRLWKPMVEE
                      MYMEEMKEQAKNMGSMEKTPLDQSNEDSASKSTSNQEKSPMADTNYHMNPNHNGDLEG
                      VTGMQGSPKRLRTSDETMMQPINADFSSNEKLTMKILEERQGIRSDGGYPFMGNFGQY
                      QMDEMSRFDVVSDQELMAQRYSGNNNGVSLTLGLPHCDSLSSTHHQGFMQTHHGIPIG
                      RRVKIGETEEYGPATINGGSSTTTAHSSAAAAAAYNGMNIQNQKRYVAQLLPDFVA"
ORIGIN

  Query Match           11.9%;  Score 324.6;  DB 8;  Length 2043;
  Best Local Similarity  65.2%;  Pred. No. 2.5e-55;
  Matches  534;  Conservative    0;  Mismatches  264;  Indels   21;  Gaps    3;

Qy        758 TTTTAGGCTCTAAGTATCTGAAAGCTGCACAAGAGCTTCTTGATGAAGTTGTTAATATTG 817
              ||  |  ||||  |||||  |||| || ||||||||||||||||  ||||| || ||   ||
Db        569 TTGTTAGCTCCAAGTACTTGAAGGCAGCACAAGAGCTTCTTGACGAAGTAGTCAACGCTG 628

Qy        818 TTGGAAAAAGCATCAAAGGAGATGATCAAAAGAAGGATAATTCAATGAATAAAGAATCAA 877
              |    |   |||  || |   |    |||                 || |  ||      ||
Db        629 ATTCCGATGACATGAACGCTAAATCCCAACTATTCTCATCGAAAAAGGGTAGTTGCGGAA 688

Qy        878 TGCCTTTGGCTAGTGATGTCAACACTAATAGTTCTGGTGGTGGTGAAAGTAGCAGCAGGC 937
                  |    ||    |  |       |      |  |||| || |  |       |  || |
Db        689 ATGATAAACCTGTCGGAGAATCATCGGCCGGCGCTGGAGGAGAAGGTTCCGGTGGCGGAG 748

Qy        938 AGAAAAATG------AAGTTGCTGTTGAGCTTACAACTGCTCAAAGACAAGAACTTCAAA 991
                ||   |      ||  | | || |||||    || ||  | |||||||||| | || |

Db        749 CAGAAGCAGCCGGGAAACGTCCGGTGGAGCTAGGCACGGCAGAGAGACAAGAAATACAGA 808

Qy        992 TGAAAAAAGCCAAGCTTCTTGCCATGCTTGAAGAGGTGGAGCAAAGGTACAGACAGTACC 1051
              |||| ||||| || |||  |  |||||||  |||||||||||  || || ||||||||||
Db        809 TGAAGAAAGCAAAACTTAGTAACATGCTTCATGAGGTGGAGCAGAGATATAGACAGTACC 868

Qy       1052 ATCACCAAATGCAAATAATTGTATTATCATTTGAGCAAGTAGCAGGAATTGGATCAGCCA 1111
              | || || ||||| ||  |  | |  || || |||||||  ||||| || |||||||| |
Db        869 ACCAGCAGATGCAGATGGTGATCTCTTCGTTCGAGCAAGCGGCAGGGATAGGATCAGCGA 928

Qy       1112 AATCATACACTCAATTAGCTTTGCATGCAATTTCGAAGCAATTCAGATGCCTAAAGGATG 1171
              | ||||||||      |||| ||| |  | || || |  || ||| | ||| | || || |
Db        929 AGTCATACACGTCGCTAGCATTGAAAACCATATCAAGACAGTTCCGTTGCTTGAAAGAGG 988

Qy       1172 CAATTGCTGAGCAAGTAAAGGCGACGAGCAAGAGTTTAGGTGAAGAGGAAGGCTTG---- 1227
              | || |||   ||  |||| ||| | | ||||||| | || || || ||     ||
Db        989 CGATCGCTGGTCAGATAAAAGCGGCCAACAAGAGTCTTGGGGAGGAAGATTCAGTGTCTG 1048

Qy       1228 --GGAGGGAAAATCGAAGGCTCAAGACTCAAATTTGTGGACCATCATCTAAGGCAACAAC 1285
                |  ||||   | || || || || ||||| || |||||||| ||  | || || |||
Db       1049 GTGTTGGGAGGTTTGAGGGGTCGAGGCTCAAGTTCGTGGACCACCACTTGAGACAGCAAA 1108

Qy       1286 GCGCGCTGCAACAGATAGGAATGATGCAAC---------CAAATGCTTGGAGACCCCAAA 1336
              | || || |||||  | |||||||| ||||           |||||||||||||| |||
Db       1109 GAGCTCTTCAACAACTGGGAATGATTCAACATCCTTCCAATAATGCTTGGAGACCTCAAC 1168

Qy       1337 GAGGTTTACCTGAAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTTTTCGAGCATTTTCTTC 1396
              | ||| | || ||| |||| ||||| || || |||||||||| ||||| || |||||||
Db       1169 GTGGTCTCCCAGAACGAGCCGTCTCAGTTCTCCGTGCTTGGCTCTTCGAACACTTTCTTC 1228

Qy       1397 ATCCTTACCCAAAGGATTCAGACAAAATCATGCTTGCTAAGCAAACGGGGCTAACAAGGA 1456
              |||| ||||| |||||||| |||||   |||||| |||||||||||| || || ||  | |
Db       1229 ATCCATACCCTAAGGATTCGGACAAGCACATGCTAGCTAAGCAAACAGGACTCACTCGTA 1288

Qy       1457 GCCAGGTGTCTAACTGGTTCATAAATGCTCGAGTTCGATTATGGAAGCCAATGGTAGAAG 1516
              |||||||||| |||||||| ||||| ||  |||||| |||||||| |||||||| || |
Db       1289 GCCAGGTGTCGAACTGGTTTATAAACGCGAGAGTTCGGTTATGGAAACCAATGGTGGAGG 1348

Qy       1517 AAATGTACTTGGAAGAAGTGAAGAATCAAGAACAAAACA 1555
              | |||||| |||| ||| ||||| | || | | ||||
Db       1349 AGATGTACATGGAGGAAATGAAGGAGCAGGCAAAGAACA 1387


RESULT 8
AY096548
LOCUS       AY096548               2074 bp    mRNA    linear   PLN 18-SEP-2002
DEFINITION  Arabidopsis thaliana putative homeodomain transcription factor
            (At2g35940) mRNA, complete cds.
ACCESSION   AY096548
VERSION     AY096548.1  GI:20465476
KEYWORDS    FLI_CDNA.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 2074)
  AUTHORS   Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M.,
            Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L.,
            Tang,C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L.,
            Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
            Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
            Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,
            Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A.,
            Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
  TITLE     Arabidopsis Open Reading Frame (ORF) Clones
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2074)
  AUTHORS   Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M.,
            Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L.,
            Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L.,
            Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
            Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
```

```
            Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,
            Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A.,
            Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-APR-2002) Plant Gene Expression Center, 800 Buchanan
            Street, Albany, CA 94710, USA
COMMENT     The RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
            Arabidopsis Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai, J.,
            Hayashizaki,Y. and Shinozaki,K.

            The Salk, Stanford, PGEC (SSP) Consortium members constructed and
            sequenced the pUNI (ORF) clones using the RAFL cDNAs: Yamada,K.,
            Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M.,
            Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C.,
            Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Chen,H.,
            Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Lam,B., Lin,J.,
            Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P.,
            Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.

            Yamada,K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally
            to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP
            /PGEC) contributed equally to this work as PIs.

            Annotation is based on the January 2002 version of the Arabidopsis
            genome submitted to GenBank.
FEATURES             Location/Qualifiers
     source          1. .2074
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /chromosome="2"
                     /clone="U18537"
                     /ecotype="Columbia"
                     /note="This clone is in pUNI 51."
     gene            1. .2074
                     /gene="At2g35940"
     CDS             1. .2043
                     /gene="At2g35940"
                     /codon_start=1
                     /evidence=experimental
                     /product="putative homeodomain transcription factor"
                     /protein_id="AAM20198.1"
                     /db_xref="GI:20465477"
                     /translation="MAAYFHGNPPEISAGSDGGLQTLILMNPTTYVQYTQQDNDSNNN
                     NNSNNSNNNNTNTNTNNNNSSFVFLDSHAPQPNASQQFVGIPLSGHEAASITAADNIS
                     VLHGYPPRVQYSLYGSHQVDPTHQQAACETPRAQQGLSLTLSSQQQQQQQHHQQHQPI
                     HVGFGSGHGEDIRVGSGSTGSGVTNGIANLVSSKYLKAAQELLDEVVNADSDDMNAKS
                     QLFSSKKGSCGNDKPVGESSAGAGGEGSGGGAEAAGKRPVELGTAERQEIQMKKAKLS
                     NMLHEVEQRYRQYHQQMQMVISSFEQAAGIGSAKSYTSLALKTISRQFRCLKEAIAGQ
                     IKAANKSLGEEDSVSGVGRFEGSRLKFVDHHLRQQRALQQLGMIQHPSNNAWRPQRGL
                     PERAVSVLRAWLFEHFLHPYPKDSDKHMLAKQTGLTRSQVSNWFINARVRLWKPMVEE
                     MYMEEMKEQAKNMGSMEKTPLDQSNEDSASKSTSNQEKSPMADTNYHMNPNHNGDLEG
                     VTGMQGSPKRLRTSDETMMQPINADFSSNEKLTMKILEERQGIRSDGGYPFMGNFGQY
                     QMDEMSRFDVVSDQELMAQRYSGNNNGVSLTLGLPHCDSLSSTHHQGFMQTHHGIPIG
                     RRVKIGETEEYGPATINGGSSTTTAHSSAAAAAAYNGMNIQNQKRYVAQLLPDFVA"
     3'UTR           2044. .2074
                     /gene="At2g35940"
ORIGIN

  Query Match           11.9%;  Score 324.6;  DB 8;  Length 2074;
  Best Local Similarity   65.2%;  Pred. No. 2.5e-55;
  Matches  534;  Conservative    0;  Mismatches  264;  Indels   21;  Gaps    3;

Qy        758 TTTTAGGCTCTAAGTATCTGAAAGCTGCACAAGAGCTTCTTGATGAAGTTGTTAATATTG 817
              || |  |||| |||||  |||| || |||||||||||||||||| ||||| || ||   ||
Db        569 TTGTTAGCTCCAAGTACTTGAAGGCAGCACAAGAGCTTCTTGACGAAGTAGTCAACGCTG 628

Qy        818 TTGGAAAAAGCATCAAAGGAGATGATCAAAAGAAGGATAATTCAATGAATAAAGAATCAA 877
              |    |   ||| || |   |    |||           || |  ||        ||
Db        629 ATTCCGATGACATGAACGCTAAATCCCAACTATTCTCATCGAAAAAGGGTAGTTGCGGAA 688

Qy        878 TGCCTTTGGCTAGTGATGTCAACACTAATAGTTCTGGTGGTGGTGAAAGTAGCAGCAGGC 937
              |     ||   | | |       |     | |||| || |  |      |  || |
Db        689 ATGATAAACCTGTCGGAGAATCATCGGCCGGCGCTGGAGGAGAAGGTTCCGGTGGCGGAG 748

Qy        938 AGAAAAATG------AAGTTGCTGTTGAGCTTACAACTGCTCAAAGACAAGAACTTCAAA 991
              ||   |      ||  | | || |||||    || ||  | |||||||| | || |
Db        749 CAGAAGCAGCCGGGAAACGTCCGGTGGAGCTAGGCACGGCAGAGAGACAAGAAATACAGA 808

Qy        992 TGAAAAAAGCCAAGCTTCTTGCCATGCTTGAAGAGGTGGAGCAAAGGTACAGACAGTACC 1051
              |||| ||||| || ||| | |||||||| | |||||||||||| || || |||||||||
Db        809 TGAAGAAAGCAAAACTTAGTAACATGCTTCATGAGGTGGAGCAGAGATATAGACAGTACC 868

Qy       1052 ATCACCAAATGCAAATAATTGTATTATCATTTGAGCAAGTAGCAGGAATTGGATCAGCCA 1111
              | || || ||||| ||  |  | |  || || |||||||  ||||| || ||||||| |
Db        869 ACCAGCAGATGCAGATGGTGATCTCTTCGTTCGAGCAAGCGGCAGGGATAGGATCAGCGA 928

Qy       1112 AATCATACACTCAATTAGCTTTGCATGCAATTTCGAAGCAATTCAGATGCCTAAAGGATG 1171
              | ||||||||     |||| ||| |  | || || |  || ||| | ||| | || || |
Db        929 AGTCATACACGTCGCTAGCATTGAAAACCATATCAAGACAGTTCCGTTGCTTGAAAGAGG 988

Qy       1172 CAATTGCTGAGCAAGTAAAGGCGACGAGCAAGAGTTTAGGTGAAGAGGAAGGCTTG---- 1227
              | || ||||  ||  |||| ||| | | ||||||| | || || || ||      ||
Db        989 CGATCGCTGGTCAGATAAAAGCGGCCAACAAGAGTCTTGGGGAGGAAGATTCAGTGTCTG 1048

Qy       1228 --GGAGGGAAAATCGAAGGCTCAAGACTCAAATTTGTGGACCATCATCTAAGGCAACAAC 1285
                |  ||||   | || || || || |||| || |||||||| ||  | || || |||
Db        1049 GTGTTGGGAGGTTTGAGGGGTCGAGGCTCAAGTTCGTGGACCACCACTTGAGACAGCAAA 1108

Qy       1286 GCGCGCTGCAACAGATAGGAATGATGCAAC---------CAAATGCTTGGAGACCCCAAA 1336
              | || || |||||  | ||||||| ||||          ||||||||||||| |||
Db        1109 GAGCTCTTCAACAACTGGGAATGATTCAACATCCTTCCAATAATGCTTGGAGACCTCAAC 1168

Qy       1337 GAGGTTTACCTGAAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTTTTCGAGCATTTTCTTC 1396
              | ||| | || ||| |||| ||||| || || |||||||||| ||||| || |||||||
Db        1169 GTGGTCTCCCAGAACGAGCCGTCTCAGTTCTCCGTGCTTGGCTCTTCGAACACTTTCTTC 1228

Qy       1397 ATCCTTACCCAAAGGATTCAGACAAAATCATGCTTGCTAAGCAAACGGGGCTAACAAGGA 1456
              |||| ||||| |||||||| |||||   |||||| |||||||||| || || ||  | |
Db        1229 ATCCATACCCTAAGGATTCGGACAAGCACATGCTAGCTAAGCAAACAGGACTCACTCGTA 1288

Qy       1457 GCCAGGTGTCTAACTGGTTCATAAATGCTCGAGTTCGATTATGGAAGCCAATGGTAGAAG 1516
              |||||||||| |||||||| ||||| ||  |||||| |||||||| |||||||| || |
Db        1289 GCCAGGTGTCGAACTGGTTTATAAACGCGAGAGTTCGGTTATGGAAACCAATGGTGGAGG 1348

Qy       1517 AAATGTACTTGGAAGAAGTGAAGAATCAAGAACAAAACA 1555
              | |||||| |||| ||| ||||| | || | | | ||||
Db        1349 AGATGTACATGGAGGAAATGAAGGAGCAGGCAAAGAACA 1387


RESULT 9
AF353094
LOCUS       AF353094                2385 bp    mRNA    linear   PLN 30-APR-2001
DEFINITION  Arabidopsis thaliana BEL1-like homeodomain 1 (BLH1) mRNA, complete
            cds.
ACCESSION   AF353094
VERSION     AF353094.1  GI:13877516
KEYWORDS    .
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 2385)
  AUTHORS   Pidkowich,M.S., Samach,A., Modrusan,Z. and Haughn,G.W.
  TITLE     A family of BEL1-like homeodomain (BLH) proteins in Arabidopsis
            thaliana
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2385)
  AUTHORS   Pidkowich,M.S., Samach,A., Modrusan,Z. and Haughn,G.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-FEB-2001) Botany, University of British Columbia,
            #3529 - 6270 University Blvd., Vancouver, BC V6T 1Z4, Canada
FEATURES             Location/Qualifiers
```

```
     source          1. .2385
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /chromosome="II"
     gene            1. .2385
                     /gene="BLH1"
     CDS             179. .2221
                     /gene="BLH1"
                     /note="homeodomain protein BLH1; transcription factor;
                     capable of heterodimerization with KNAT proteins; similar
                     to predicted homeodomain protein presented in GenBank
                     Accession Number AAD21463"
                     /codon_start=1
                     /product="BEL1-like homeodomain 1"
                     /protein_id="AAK43836.1"
                     /db_xref="GI:13877517"
                     /translation="MAAYFHGNPPEISAGSDGGLQTLILMNPTTYVQYTQQDNDSNNN
                     NNSNNSNNNNTNTNTNNNNSSFVFLDSHAPQPNASQQFVGIPLSGHEAASITAADNIS
                     VLHGYPPRVQYSLYGSHQVDPTHQQAACETPRAQQGLSLTLSSQQQQQQQHHQQHQPI
                     HVGFGSGHGEDIRVGSGSTGSGVTNGIANLVSSKYLKAAQELLDEVVNADSDDMNAKS
                     QLFSSKKGSCGNDKPVGESSAGAGGEGSGGGAEAAGKRPVELGTAERQEIQMKKAKLS
                     NMLHEVEQRYRQYHQQMQMVISSFEQAAGIGSAKSYTSLALKTISRQFRCLKEAIAGQ
                     IKAANKSLGEEDSVSGVGRFEGSRLKFVDHHLRQQRALQQLGMIQHPSNNAWRPQRGL
                     PERAVSVLRAWLFEHFLHPYPKDSDKHMLAKQTGLTRSQVSNWFINARVRLWKPMVEE
                     MYMEEMKEQAKNMGSMEKTPLDQSNEDSASKSTSNQEKSPMADTNYHMNPNHNGDLEG
                     VTGMQGCPKRLRTSDETMMQPINADFSSNEKLTMKILEERQGIRSDGGYPFMGNFGQY
                     QMDEMSRFDVVSDQELMAQRYSGNNNGVSLTLGLPHCDSLSSTDHQGFMQTHHGIPIG
                     RRVKIGETEEYGPATINGGSSTTTAHSSAAAAAAYNGMNIQNQKRYVAQLLPDFVA"
ORIGIN

  Query Match           11.9%;  Score 324.6;  DB 8;  Length 2385;
  Best Local Similarity  65.2%;  Pred. No. 2.5e-55;
  Matches  534;  Conservative    0;  Mismatches  264;  Indels   21;  Gaps    3;

Qy        758 TTTTAGGCTCTAAGTATCTGAAAGCTGCACAAGAGCTTCTTGATGAAGTTGTTAATATTG 817
              ||  |  ||||  |||||  |||| ||  ||||||||||||||||||  |||||  ||  ||    ||
Db        747 TTGTTAGCTCCAAGTACTTGAAGGCAGCACAAGAGCTTCTTGACGAAGTAGTCAACGCTG 806

Qy        818 TTGGAAAAAGCATCAAAGGAGATGATCAAAAGAAGGATAATTCAATGAATAAAGAATCAA 877
              |    |   |||  ||  |    |    |||                  ||  |   ||        ||
Db        807 ATTCCGATGACATGAACGCTAAATCCCAACTATTCTCATCGAAAAAGGGTAGTTGCGGAA 866

Qy        878 TGCCTTTGGCTAGTGATGTCAACACTAATAGTTCTGGTGGTGGTGAAAGTAGCAGCAGGC 937
                  |    ||    |  |       |     |  ||||  ||  |  |       |  |||
Db        867 ATGATAAACCTGTCGGAGAATCATCGGCCGGCGCTGGAGGAGAAGGTTCCGGTGGCGGAG 926

Qy        938 AGAAAAATG------AAGTTGCTGTTGAGCTTACAACTGCTCAAAGACAAGAACTTCAAA 991
                ||    |      ||  |  | ||  ||||     ||  ||   |  |||||||||  |  ||  |
Db        927 CAGAAGCAGCCGGGAAACGTCCGGTGGAGCTAGGCACGGCAGAGAGACAAGAAATACAGA 986

Qy        992 TGAAAAAAGCCAAGCTTCTTGCCATGCTTGAAGAGGTGGAGCAAAGGTACAGACAGTACC 1051
              ||||  |||||  ||  |||   |  |||||||  |  ||||||||||  ||  ||  ||||||||||
Db        987 TGAAGAAAGCAAAACTTAGTAACATGCTTCATGAGGTGGAGCAGAGATATAGACAGTACC 1046

Qy       1052 ATCACCAAATGCAAATAATTGTATTATCATTTGAGCAAGTAGCAGGAATTGGATCAGCCA 1111
              |  ||  ||  |||||  ||   |  |  ||  ||  |||||||   |||||  ||  ||||||||  |
Db       1047 ACCAGCAGATGCAGATGGTGATCTCTTCGTTCGAGCAAGCGGCAGGGATAGGATCAGCGA 1106

Qy       1112 AATCATACACTCAATTAGCTTTGCATGCAATTTCGAAGCAATTCAGATGCCTAAAGGATG 1171
              |  ||||||||      ||||  |||  |   |  ||  ||  |   ||  |||  |  ||||  |  ||  |
Db       1107 AGTCATACACGTCGCTAGCATTGAAAACCATATCAAGACAGTTCCGTTGCTTGAAAGAGG 1166

Qy       1172 CAATTGCTGAGCAAGTAAAGGCGACGAGCAAGAGTTTAGGTGAAGAGGAAGGCTTG---- 1227
              |  ||  ||||   ||   ||||  |||  |  |  |||||||  |  ||  ||  ||  ||      ||
Db       1167 CGATCGCTGGTCAGATAAAAGCGGCCAACAAGAGTCTTGGGGAGGAAGATTCAGTGTCTG 1226

Qy       1228 --GGAGGGAAAATCGAAGGCTCAAGACTCAAATTTGTGGACCATCATCTAAGGCAACAAC 1285
                |  ||||    |  ||  ||  ||  ||  |||||  ||  ||||||||  ||   |  ||  ||  |||
Db       1227 GTGTTGGGAGGTTTGAGGGGTCGAGGCTCAAGTTCGTGGACCACCACTTGAGACAGCAAA 1286

Qy       1286 GCGCGCTGCAACAGATAGGAATGATGCAAC---------CAAATGCTTGGAGACCCCAAA 1336
              |  ||  ||  |||||   |  ||||||||  ||||           ||||||||||||||  |||
```

```
Db       1287 GAGCTCTTCAACAACTGGGAATGATTCAACATCCTTCCAATAATGCTTGGAGACCTCAAC 1346

Qy       1337 GAGGTTTACCTGAAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTTTTCGAGCATTTTCTTC 1396
              |  |||  |  ||  |||  ||||  |||||  ||  ||  ||||||||||  |||||  ||  |||||||
Db       1347 GTGGTCTCCCAGAACGAGCCGTCTCAGTTCTCCGTGCTTGGCTCTTCGAACACTTTCTTC 1406

Qy       1397 ATCCTTACCCAAAGGATTCAGACAAAATCATGCTTGCTAAGCAAACGGGGCTAACAAGGA 1456
              ||||  |||||  ||||||||  |||||    ||||||  ||||||||||  ||  ||  ||    |  |
Db       1407 ATCCATACCCTAAGGATTCGGACAAGCACATGCTAGCTAAGCAAACAGGACTCACTCGTA 1466

Qy       1457 GCCAGGTGTCTAACTGGTTCATAAATGCTCGAGTTCGATTATGGAAGCCAATGGTAGAAG 1516
              ||||||||||  ||||||||  |||||  ||    |||||||  ||||||||  ||||||||  ||  |
Db       1467 GCCAGGTGTCGAACTGGTTTATAAACGCGAGAGTTCGGTTATGGAAACCAATGGTGGAGG 1526

Qy       1517 AAATGTACTTGGAAGAAGTGAAGAATCAAGAACAAAACA 1555
              |  ||||||  ||||  |||  |||||  |  ||  |  |  |  ||||
Db       1527 AGATGTACATGGAGGAAATGAAGGAGCAGGCAAAGAACA 1565
```

```
RESULT 10
AY072175
LOCUS       AY072175               2572 bp    mRNA    linear   PLN 18-SEP-2002
DEFINITION  Arabidopsis thaliana putative homeodomain transcription factor
            (At2g35940) mRNA, complete cds.
ACCESSION   AY072175
VERSION     AY072175.1  GI:18176172
KEYWORDS    FLI_CDNA.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 2572)
  AUTHORS   Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K.,
            Goldsmith,A.D., Lee,J.M., Quach,H.L., Toriumi,M., Yu,G., Bowser,L.,
            Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
            Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
            Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T.,
            Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K.,
            Davis,R.W., Ecker,J.R. and Theologis,A.
  TITLE     Arabidopsis Full Length cDNA Clones
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2572)
  AUTHORS   Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M.,
            Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L.,
            Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L.,
            Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
            Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
            Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,
            Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A.,
            Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JAN-2002) Plant Gene Expression Center, 800 Buchanan
            Street, Albany, CA 94710, USA
COMMENT     RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
            Arabidopsis Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
            Hayashizaki,Y. and Shinozaki,K.

            The Salk, Stanford, PGEC (SSP) Consortium members carried out the
            sequencing and annotation of the RAFL cDNAs: Yamada,K., Banh,J.,
            Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M.,
            Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C.,
            Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Chen,H.,
            Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Lam,B., Lin,J.,
            Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P.,
            Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.

            Yamada,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally to
            this work.  Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC)
            contributed equally to this work as PIs.
```

```
          Annotation is based on the January 2002 version of the Arabidopsis
          genome submitted to GenBank.
FEATURES             Location/Qualifiers
     source          1. .2572
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /chromosome="2"
                     /clone="RAFL09-25-N04 (R18537)"
                     /ecotype="Columbia"
                     /note="This clone is in a modified pBluescript vector
                     (FLC-1) as a BamHI/XhoI insert."
     gene            1. .2572
                     /gene="At2g35940"
     5'UTR           1. .205
                     /gene="At2g35940"
     misc_difference 2
                     /gene="At2g35940"
                     /note="compared to genomic sequence"
                     /replace="c"
     CDS             206. .2248
                     /gene="At2g35940"
                     /codon_start=1
                     /evidence=experimental
                     /product="putative homeodomain transcription factor"
                     /protein_id="AAL59997.1"
                     /db_xref="GI:18176173"
                     /translation="MAAYFHGNPPEISAGSDGGLQTLILMNPTTYVQYTQQDNDSNNN
                     NNSNNSNNNNTNTNTNNNNSSFVFLDSHAPQPNASQQFVGIPLSGHEAASITAADNIS
                     VLHGYPPRVQYSLYGSHQVDPTHQQAACETPRAQQGLSLTLSSQQQQQQQHHQQHQPI
                     HVGFGSGHGEDIRVGSGSTGSGVTNGIANLVSSKYLKAAQELLDEVVNADSDDMNAKS
                     QLPSSKKGSCGNDKPVGESSAGAGGEGSGGGAEAAGKRPVELGTAERQEIQMKKAKLS
                     NMLHEVEQRYRQYHQQMQMVISSFEQAAGIGSAKSYTSLALKTISRQFRCLKEAIAGQ
                     IKAANKSLGEEDSVSGVGRFEGSRLKFVDHHLRQQRALQQLGMIQHPSNNAWRPQRGL
                     PERAVSVLRAWLFEHFLHPYPKDSDKHMLAKQTGLTRSQVSNWFINARVRLWKPMVEE
                     MYMEEMKEQAKNMGSMEKTPLDQSNEDSASKSTSNQEKSPMADTNYHMNPNHNGDLEG
                     VTGMQGSPKRLRTSDETMMQPINADFSSNEKLTMKILEERQGIRSDGGYPFMGNFGQY
                     QMDEMSRFDVVSDQELMAQRYSGNNNGVSLTLGLPHCDSLSSTHHQGFMQTHHGIPIG
                     RRVKIGETEEYGPATINGGSSTTTAHSSAAAAAAYNGMNIQNQKRYVAQLLPDFVA"
     3'UTR           2249. .2572
                     /gene="At2g35940"
     misc_difference 2417. .2418
                     /gene="At2g35940"
                     /note="compared to genomic sequence"
                     /replace="ttc"
     misc_difference 2556
                     /gene="At2g35940"
                     /note="not present in genomic sequence"
ORIGIN

  Query Match            11.9%;  Score 324.6;  DB 8;  Length 2572;
  Best Local Similarity  65.2%;  Pred. No. 2.5e-55;
  Matches  534;  Conservative    0;  Mismatches  264;  Indels   21;  Gaps    3;

Qy        758 TTTTAGGCTCTAAGTATCTGAAAGCTGCACAAGAGCTTCTTGATGAAGTTGTTAATATTG 817
              || |  |||| |||||  |||| || ||||||||||||||||||| ||||| || ||   ||
Db        774 TTGTTAGCTCCAAGTACTTGAAGGCAGCACAAGAGCTTCTTGACGAAGTAGTCAACGCTG 833

Qy        818 TTGGAAAAAGCATCAAAGGAGATGATCAAAAGAAGGATAATTCAATGAATAAAGAATCAA 877
                |   |  ||| || |   |   |||           || |  ||       ||
Db        834 ATTCCGATGACATGAACGCTAAATCCCAACTATTCTCATCGAAAAAGGGTAGTTGCGGAA 893

Qy        878 TGCCTTTGGCTAGTGATGTCAACACTAATAGTTCTGGTGGTGGTGAAAGTAGCAGCAGGC 937
                  |   ||   | |       |     | |||| || |  |      |  || |
Db        894 ATGATAAACCTGTCGGAGAATCATCGGCCGGCGCTGGAGGAGAAGGTTCCGGTGGCGGAG 953

Qy        938 AGAAAAATG------AAGTTGCTGTTGAGCTTACAACTGCTCAAAGACAAGAACTTCAAA 991
                 ||   |      ||  | | || |||||    || ||  | |||||||| | || |
Db        954 CAGAAGCAGCCGGGAAACGTCCGGTGGAGCTAGGCACGGCAGAGAGACAAGAAATACAGA 1013

Qy        992 TGAAAAAAGCCAAGCTTCTTGCCATGCTTGAAGAGGTGGAGCAAAGGTACAGACAGTACC 1051
              |||| ||||| || |||  |  ||||||| | |||||||||||| || || ||||||||||
Db       1014 TGAAGAAAGCAAAACTTAGTAACATGCTTCATGAGGTGGAGCAGAGATATAGACAGTACC 1073

Qy       1052 ATCACCAAATGCAAATAATTGTATTATCATTTGAGCAAGTAGCAGGAATTGGATCAGCCA 1111
              | || || ||||| ||  |  | |  || || |||||||  ||||| || ||||||||  |
Db       1074 ACCAGCAGATGCAGATGGTGATCTCTTCGTTCGAGCAAGCGGCAGGGATAGGATCAGCGA 1133

Qy       1112 AATCATACACTCAATTAGCTTTGCATGCAATTTCGAAGCAATTCAGATGCCTAAAGGATG 1171
              | ||||||||     |||| ||| |  | || || |  || ||| | ||| | || || |
Db       1134 AGTCATACACGTCGCTAGCATTGAAAACCATATCAAGACAGTTCCGTTGCTTGAAAGAGG 1193

Qy       1172 CAATTGCTGAGCAAGTAAAGGCGACGAGCAAGAGTTTAGGTGAAGAGGAAGGCTTG---- 1227
              | || ||||   ||  |||| ||| | | |||||||  | || || || ||      ||
Db       1194 CGATCGCTGGTCAGATAAAAGCGGCCAACAAGAGTCTTGGGGAGGAAGATTCAGTGTCTG 1253

Qy       1228 --GGAGGGAAAATCGAAGGCTCAAGACTCAAATTTGTGGACCATCATCTAAGGCAACAAC 1285
                |  ||||   | || || || || ||||| || |||||||| ||  | || || |||
Db       1254 GTGTTGGGAGGTTTGAGGGGTCGAGGCTCAAGTTCGTGGACCACCACTTGAGACAGCAAA 1313

Qy       1286 GCGCGCTGCAACAGATAGGAATGATGCAAC---------CAAATGCTTGGAGACCCCAAA 1336
              | || || |||||  | |||||||| ||||          |||||||||||||| |||
Db       1314 GAGCTCTTCAACAACTGGGAATGATTCAACATCCTTCCAATAATGCTTGGAGACCTCAAC 1373

Qy       1337 GAGGTTTACCTGAAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTTTTCGAGCATTTTCTTC 1396
              | ||| | || ||| |||| ||||| || || |||||||||||| ||||| || |||||||
Db       1374 GTGGTCTCCCAGAACGAGCCGTCTCAGTTCTCCGTGCTTGGCTCTTCGAACACTTTCTTC 1433

Qy       1397 ATCCTTACCCAAAGGATTCAGACAAAATCATGCTTGCTAAGCAAACGGGGCTAACAAGGA 1456
              |||| ||||| |||||||| |||||   |||||| |||||||||| || || ||  | |
Db       1434 ATCCATACCCTAAGGATTCGGACAAGCACATGCTAGCTAAGCAAACAGGACTCACTCGTA 1493

Qy       1457 GCCAGGTGTCTAACTGGTTCATAAATGCTCGAGTTCGATTATGGAAGCCAATGGTAGAAG 1516
              |||||||||| ||||||||  ||||| ||  |||||||  |||||||| |||||||| || |
Db       1494 GCCAGGTGTCGAACTGGTTTATAAACGCGAGAGTTCGGTTATGGAAACCAATGGTGGAGG 1553

Qy       1517 AAATGTACTTGGAAGAAGTGAAGAATCAAGAACAAAACA 1555
              | |||||| ||| ||| ||||| | || | | | ||||
Db       1554 AGATGTACATGGAGGAAATGAAGGAGCAGGCAAAGAACA 1592


RESULT 11
AY099854
LOCUS       AY099854               2578 bp    mRNA    linear   PLN 06-MAY-2002
DEFINITION  Arabidopsis thaliana putative homeodomain transcription factor
            (At2g35940) mRNA, complete cds.
ACCESSION   AY099854
VERSION     AY099854.1  GI:20466775
KEYWORDS    FLI_CDNA.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 2578)
  AUTHORS   Nguyen,M., Karlin-Neumann,G., Southwick,A., Lam,B., Miranda,M.,
            Palm,C.J., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H.,
            Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A.,
            Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K.,
            Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K.,
            Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-APR-2002) DNA Sequencing and Technology Center,
            Stanford University, 855 California Avenue, Palo Alto, CA 94304,
            USA
COMMENT     e-mail for correspondence: arab@sequence.stanford.edu

            RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
            Arabidopsis  Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
            Hayashizaki,Y. and Shinozaki,K.

            The Salk, Stanford, PGEC (SSP) Consortium members carried out the
            sequencing and annotation of the RAFL cDNAs: Nguyen,M.,
            Southwick,A., Karlin-Neumann,G.,  Lam,B., Miranda,M., Palm,C.J.,
```

Bowser,L., Jones,T., Banh,J., Chen,H., Cheuk,R., Chung,M.K., Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H., Shinn,P., Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.

Nguyen,M, (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W. (SSP/Stanford) contributed equally to this work as PIs.

FEATURES Location/Qualifiers
source 1. .2578
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="2"
/clone="RAFL09-94-K12 (R25364)"
/ecotype="Columbia"
/note="This clone is in pBluescript"
gene 1. .2578
/gene="At2g35940"
/note="synonym: F11F19.15"
CDS 198. .2240
/gene="At2g35940"
/codon_start=1
/product="putative homeodomain transcription factor"
/protein_id="AAM20705.1"
/db_xref="GI:20466776"
/translation="MAAYFHGNPPEISAGSDGGLQTLILMNPTTYVQYTQQDNDSNNN
NNSNNSNNNNTNTNTNNNNSSFVFLDSHAPQPNASQQFVGIPLSGHEAASITAADNIS
VLHGYPPRVQYSLYGSHQVDPTHQQAACETPRAQQGLSLTLSSQQQQQQQHHQQHQPI
HVGFGSGHGEDIRVGSGSTGSGVTNGIANLVSSKYLKAAQELLDEVVNADSDDMNAKS
QLFSSKKGSCGNDKPVGESSAGAGGEGSGGGAEAAGKRPVELGTAERQEIQMKKAKLS
NMLHEVEQRYRQYHQQMQMVISSFEQAAGIGSAKSYTSLALKTISRQFRCLKEAIAGQ
IKAANKSLGEEDSVSGVGRFEGSRLKFVDHHLRQQRALQQLGMIQHPSNNAWRPQRGL
PERAVSVLRAWLFEHFLHPYPKDSDKHMLAKQTGLTRSQVSNWFINARVRLWKPMVEE
MYMEEMKEQAKNMGSMEKTPLDQSNEDSASKSTSNQEKSPMADTNYHMNPNHNGDLEG
VTGMQGSPKRLRTSDETMMQPINADFSSNEKLTMKILEERQGIRSDGGYPFMGNFGQY
QMDEMSRFDVVSDQELMAQRYSGNNNGVSLTLGLPHCDSLSSTHHQGFMQTHHGIPIG
RRVKIGETEEYGPATINGGSSTTTAHSSAAAAAAYNGMNIQNQKRYVAQLLPDFVA"

ORIGIN

Query Match 11.9%; Score 324.6; DB 8; Length 2578;
Best Local Similarity 65.2%; Pred. No. 2.5e-55;
Matches 534; Conservative 0; Mismatches 264; Indels 21; Gaps 3;

```
Qy       758 TTTTAGGCTCTAAGTATCTGAAAGCTGCACAAGAGCTTCTTGATGAAGTTGTTAATATTG 817
Db       766 TTGTTAGCTCCAAGTACTTGAAGGCAGCACAAGAGCTTCTTGACGAAGTAGTCAACGCTG 825

Qy       818 TTGGAAAAAGCATCAAAGGAGATGATCAAAAGAAGGATAATTCAATGAATAAAGAATCAA 877
Db       826 ATTCCGATGACATGAACGCTAAATCCCAACTATTCTCATCGAAAAAGGGTAGTTGCGGAA 885

Qy       878 TGCCTTTGGCTAGTGATGTCAACACTAATAGTTCTGGTGGTGGTGAAAGTAGCAGCAGGC 937
Db       886 ATGATAAACCTGTCGGAGAATCATCGGCCGGCGCTGGAGGAGAAGGTTCCGGTGGCGGAG 945

Qy       938 AGAAAAATG------AAGTTGCTGTTGAGCTTACAACTGCTCAAAGACAAGAACTTCAAA 991
Db       946 CAGAAGCAGCCGGGAAACGTCCGGTGGAGCTAGGCACGGCAGAGAGACAAGAAATACAGA 1005

Qy       992 TGAAAAAAGCCAAGCTTCTTGCCATGCTTGAAGAGGTGGAGCAAAGGTACAGACAGTACC 1051
Db      1006 TGAAGAAAGCAAAACTTAGTAACATGCTTCATGAGGTGGAGCAGAGATATAGACAGTACC 1065

Qy      1052 ATCACCAAATGCAAATAATTGTATTATCATTTGAGCAAGTAGCAGGAATTGGATCAGCCA 1111
Db      1066 ACCAGCAGATGCAGATGGTGATCTCTTCGTTCGAGCAAGCGGCAGGGATAGGATCAGCGA 1125

Qy      1112 AATCATACACTCAATTAGCTTTGCATGCAATTTCGAAGCAATTCAGATGCCTAAAGGATG 1171
Db      1126 AGTCATACACGTCGCTAGCATTGAAAACCATATCAAGACAGTTCCGTTGCTTGAAAGAGG 1185

Qy      1172 CAATTGCTGAGCAAGTAAAGGCGACGAGCAAGAGTTTAGGTGAAGAGGAAGGCTTG---- 1227
```

```
Db      1186 CGATCGCTGGTCAGATAAAAGCGGCCAACAAGAGTCTTGGGGAGGAAGATTCAGTGTCTG 1245

Qy      1228 --GGAGGGAAAATCGAAGGCTCAAGACTCAAATTTGTGGACCATCATCTAAGGCAACAAC 1285
Db      1246 GTGTTGGGAGGTTTGAGGGGTCGAGGCTCAAGTTCGTGGACCACCACTTGAGACAGCAAA 1305

Qy      1286 GCGCGCTGCAACAGATAGGAATGATGCAAC---------CAAATGCTTGGAGACCCCAAA 1336
Db      1306 GAGCTCTTCAACAACTGGGAATGATTCAACATCCTTCCAATAATGCTTGGAGACCTCAAC 1365

Qy      1337 GAGGTTTACCTGAAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTTTTCGAGCATTTTCTTC 1396
Db      1366 GTGGTCTCCCAGAACGAGCCGTCTCAGTTCTCCGTGCTTGGCTCTTCGAACACTTTCTTC 1425

Qy      1397 ATCCTTACCCAAAGGATTCAGACAAAATCATGCTTGCTAAGCAAACGGGGCTAACAAGGA 1456
Db      1426 ATCCATACCCTAAGGATTCGGACAAGCACATGCTAGCTAAGCAAACAGGACTCACTCGTA 1485

Qy      1457 GCCAGGTGTCTAACTGGTTCATAAATGCTCGAGTTCGATTATGGAAGCCAATGGTAGAAG 1516
Db      1486 GCCAGGTGTCGAACTGGTTTATAAACGCGAGAGTTCGGTTATGGAAACCAATGGTGGAGG 1545

Qy      1517 AAATGTACTTGGAAGAAGTGAAGAATCAAGAACAAAACA 1555
Db      1546 AGATGTACATGGAGGAAATGAAGGAGCAGGCAAAGAACA 1584
```

RESULT 12
AK104872
LOCUS AK104872 1896 bp mRNA linear PLN 24-JUL-2003
DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:001-044-C01, full insert sequence.
ACCESSION AK104872
VERSION AK104872.1 GI:32990081
KEYWORDS FLI_CDNA; oligo capping.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1
AUTHORS The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team:, Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group:, Otomo,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y., Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M., Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J., Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J., Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN:, Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y., Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Yoshino,M. and Hayashizaki,Y.
TITLE Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice
JOURNAL Science 301 (5631), 376-379 (2003)
MEDLINE 22752273
PUBMED 12869764
REFERENCE 2 (bases 1 to 1896)
AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K., Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K., Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M., Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M., Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M.,

Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A., Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M., Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K., Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H., Osato,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K., Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S., Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W., Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and Yoshimura,A.

TITLE Direct Submission

JOURNAL Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)

COMMENT This clone is one of the 28K full-length cDNA clones from japonica rice.
URL : http://cdna01.dna.affrc.go.jp/cDNA/
NIAS Rice Full-Length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and Yamamoto,M.
FAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y., Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M., Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J., Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S., Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S., Yoshimura,A., Matsubara,K. and Murakami,K.
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J., Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N., Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K., Yasunishi,A. and Hayashizaki,Y.

FEATURES Location/Qualifiers
source 1. .1896
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="001-044-C01"

ORIGIN

Query Match 10.8%; Score 295.6; DB 8; Length 1896;
Best Local Similarity 60.3%; Pred. No. 1.9e-49;
Matches 528; Conservative 0; Mismatches 339; Indels 9; Gaps 2;

```
Qy        765 CTCTAAGTATCTGAAAGCTGCACAAGAGCTTCTTGATGAAGTTGTTAATATTGTTGGAAA 824
Db        301 CTCAAGGTATCTCAAGGCAGCACAAGAACTGCTTGATGAGGTCGTGAGTGTTTGGAAGAG 360

Qy        825 AAGCATCAAAGGAGATGATCAAAAGAAGGATAATTCAATGAATAAAGAATCAATGCCTTT 884
Db        361 CATAAAACAAAAGGCTCAGAAAGAGAAGGTTGAATCAGGAAAAGCAGATGGCAAA----- 415

Qy        885 GGCTAGTGATGTCAACACTAATAGTTCTGGTGGTGGTGAAAGTAGCAGCAGGCAGAAAAA 944
Db        416 -GAGACTGATGGAGGGCCAAAAAGTGAGGGTGTATCTTCTAATCCACAAGAATCTGGCGC 474

Qy        945 TGAAGTTGCTGTTGAGCTTACAACTGCTCAAAGACAAGAACTTCAAATGAAAAAAGCCAA 1004
```

```
Db        475 TAATGCAGCACCTGAGCTTTCCACTGCTGAGAAGCAAGAGCTTCAGAACAAAATGGCAAA 534

Qy       1005 GCTTCTTGCCATGCTTGAAGAGGTGGAGCAAAGGTACAGACAGTACCATCACCAAATGCA 1064
Db        535 ACTGATGGCAATGTTGGATGAGGTGGACCGGAAATACAAGCATTATTACCACCAAATGCA 594

Qy       1065 AATAATTGTATTATCATTTGAGCAAGTAGCAGGAATTGGATCAGCCAAATCATACACTCA 1124
Db        595 AACCGTGGTTTCATCGTTTGATGTGGTAGCTGGACCAGGATCTGCGAAGCCTTACACAGC 654

Qy       1125 ATTAGCTTTGCATGCAATTTCGAAGCAATTCAGATGCCTAAAGGATGCAATTGCTGAGCA 1184
Db        655 AGTTGCTCTTCAGACAATCTCACGGCACTTTCGGTGCCTGAAGGATGCTATCAATGATCA 714

Qy       1185 AGTAAAGGCGACGAGCAAGAGTTTAGGTGAAGAGGAAGGCTTGGGAGGGAAAATCGAAGG 1244
Db        715 GATCAATGTTATCAGGAAAAAGCTTGGCGAGGAAGAGAACTCATCTGGCAAGGAGGGAAA 774

Qy       1245 CTCAA---GACTCAAATTTGTGGACCATCATCTAAGGCAACAACGCGCGCTGCAACAGAT 1301
Db        775 ATTAACCCGCCTCCGGTACATTGATCAGCAGCTAAGGCAACAACGTGCTTTCCAACAGTA 834

Qy       1302 AGGAATGATGCAACCAAATGCTTGGAGACCCCAAAGAGGTTTACCTGAAAGAGCTGTCTC 1361
Db        835 CGGTATGATTCCACAAAACGCCTGGAGACCACAGAGGGGACTGCCTGAAAACTCTGTTAC 894

Qy       1362 TGTCCTTCGTGCTTGGCTTTTCGAGCATTTTCTTCATCCTTACCCAAAGGATTCAGACAA 1421
Db        895 AATTCTTCGCGCTTGGCTGTTTGAACACTTCCTCCATCCGTACCCAAAAGATTCCGAGAA 954

Qy       1422 AATCATGCTTGCTAAGCAAACGGGGCTAACAAGGAGCCAGGTGTCTAACTGGTTCATAAA 1481
Db        955 GTTAATGCTTGCTAGGCAGACTGGCCTGACTAGGAGTCAGATTTCGAATTGGTTCATAAA 1014

Qy       1482 TGCTCGAGTTCGATTATGGAAGCCAATGGTAGAAGAAATGTACTTGGAAGAAGTGAAGAA 1541
Db       1015 TGCCCGTGTCCGCCTTTGGAAACCAATGATTGAAGACATGTACAAAGAAGAGATTGGTGA 1074

Qy       1542 TCAAGAACAAAACAGTACTAATACTTCAGGAGATAACAAAAACAAAGAGACCAATATAAG 1601
Db       1075 TTTAGAGCAAGACTCCAACTCTTCCTCTGACAATGCACCAAGAAGTAAGGATAAAATGGC 1134

Qy       1602 TGCTCCAAATGAAGAGAAACATCCAATTATTACTAG 1637
Db       1135 ATCTTCAGAAGACAAGGAGGATCTCAAAAACTCCAG 1170
```

RESULT 13
AK121356
LOCUS AK121356 2476 bp mRNA linear PLN 29-OCT-2003
DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:J023122O08, full insert sequence.
ACCESSION AK121356
VERSION AK121356.1 GI:37990979
KEYWORDS FLI_CDNA; CAP trapper.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1
AUTHORS The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team:, Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group:, Otomo,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y., Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M., Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J., Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J., Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN:,

Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y., Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Yoshino,M. and Hayashizaki,Y.

TITLE Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice

JOURNAL Science 301 (5631), 376-379 (2003)

MEDLINE 22752273

PUBMED 12869764

REFERENCE 2

AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K., Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K., Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M., Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M., Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M., Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A., Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakahama,Y., Nakamura,M., Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K., Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H., Osato,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K., Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S., Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W., Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and Yoshimura,A.

TITLE Collection, mapping, and annotation of 28K full-length cDNA clones from japonica rice

JOURNAL Unpublished

REFERENCE 3 (bases 1 to 2476)

AUTHORS Kikuchi,S.

TITLE Direct Submission

JOURNAL Submitted (31-JAN-2003) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)

COMMENT This clone is one of the 32K full-length cDNA clones from japonica rice.
URL : http://cdna01.dna.affrc.go.jp/cDNA/
NIAS Rice Full-Length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., Yamamoto,M. and Nakahama,Y.
FAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y., Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M., Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J., Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S., Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S., Yoshimura,A., Matsubara,K. and Murakami,K.
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J., Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N., Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K., Yasunishi,A. and Hayashizaki,Y.

FEATURES Location/Qualifiers

source 1. .2476
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J023122008"

ORIGIN

Query Match 10.8%; Score 295.6; DB 8; Length 2476;
Best Local Similarity 60.3%; Pred. No. 1.9e-49;
Matches 528; Conservative 0; Mismatches 339; Indels 9; Gaps 2;

```
Qy       765 CTCTAAGTATCTGAAAGCTGCACAAGAGCTTCTTGATGAAGTTGTTAATATTGTTGGAAA 824
            ||| | |||||| || || ||||||||| || |||||||| || || | | ||       |
Db      1339 CTCAAGGTATCTCAAGGCAGCACAAGAACTGCTTGATGAGGTCGTGAGTGTTTGGAAGAG 1398

Qy       825 AAGCATCAAAGGAGATGATCAAAAGAAGGATAATTCAATGAATAAAGAATCAATGCCTTT 884
             |  |   ||   | | |  || |||||| | | |||   ||   |||    |
Db      1399 CATAAAACAAAAGGCTCAGAAAGAGAAGGTTGAATCAGGAAAAGCAGATGGCAAA----- 1453

Qy       885 GGCTAGTGATGTCAACACTAATAGTTCTGGTGGTGGTGAAAGTAGCAGCAGGCAGAAAAA 944
             | | |||||      | || |||    ||||    |   | |
Db      1454 -GAGACTGATGGAGGGCCAAAAAGTGAGGGTGTATCTTCTAATCCACAAGAATCTGGCGC 1512

Qy       945 TGAAGTTGCTGTTGAGCTTACAACTGCTCAAAGACAAGAACTTCAAATGAAAAAAGCCAA 1004
             | | |  ||   |||||| | |||||| | |  ||||| ||||| |  ||||  || ||
Db      1513 TAATGCAGCACCTGAGCTTTCCACTGCTGAGAAGCAAGAGCTTCAGAACAAAATGGCAAA 1572

Qy      1005 GCTTCTTGCCATGCTTGAAGAGGTGGAGCAAAGGTACAGACAGTACCATCACCAAATGCA 1064
             ||  | || ||| | || |||||||| |   | ||||  || ||  | ||||||||||||
Db      1573 ACTGATGGCAATGTTGGATGAGGTGGACCGGAAATACAAGCATTATTACCACCAAATGCA 1632

Qy      1065 AATAATTGTATTATCATTTGAGCAAGTAGCAGGAATTGGATCAGCCAAATCATACACTCA 1124
             ||   | || | ||| |||||     ||||| |||   ||||| || ||  | |||||
Db      1633 AACCGTGGTTTCATCGTTTGATGTGGTAGCTGGACCAGGATCTGCGAAGCCTTACACAGC 1692

Qy      1125 ATTAGCTTTGCATGCAATTTCGAAGCAATTCAGATGCCTAAAGGATGCAATTGCTGAGCA 1184
             | | ||| | ||  |||| ||   ||| ||  | ||||| |||||||| ||   ||| ||
Db      1693 AGTTGCTCTTCAGACAATCTCACGGCACTTTCGGTGCCTGAAGGATGCTATCAATGATCA 1752

Qy      1185 AGTAAAGGCGACGAGCAAGAGTTTAGGTGAAGAGGAAGGCTTGGGAGGGAAAATCGAAGG 1244
             | || |  |  || || |  | || || || ||   ||      || ||    | |
Db      1753 GATCAATGTTATCAGGAAAAAGCTTGGCGAGGAAGAGAACTCATCTGGCAAGGAGGGAAA 1812

Qy      1245 CTCAA---GACTCAAATTTGTGGACCATCATCTAAGGCAACAACGCGCGCTGCAACAGAT 1301
             | ||   | |||    |  | || || || ||||||||||||||| ||  | ||||||
Db      1813 ATTAACCCGCCTCCGGTACATTGATCAGCAGCTAAGGCAACAACGTGCTTTCCAACAGTA 1872

Qy      1302 AGGAATGATGCAACCAAATGCTTGGAGACCCCAAAGAGGTTTACCTGAAAGAGCTGTCTC 1361
             || ||||| | || ||| || |||||||| || || ||  | |||||||   ||||  |
Db      1873 CGGTATGATTCCACAAAACGCCTGGAGACCACAGAGGGGACTGCCTGAAAACTCTGTTAC 1932

Qy      1362 TGTCCTTCGTGCTTGGCTTTTCGAGCATTTTCTTCATCCTTACCCAAAGGATTCAGACAA 1421
             | ||||| |||||||| || || || || || ||||| |||||||| ||||| || ||
Db      1933 AATTCTTCGCGCTTGGCTGTTTGAACACTTCCTCCATCCGTACCCAAAAGATTCCGAGAA 1992

Qy      1422 AATCATGCTTGCTAAGCAAACGGGGCTAACAAGGAGCCAGGTGTCTAACTGGTTCATAAA 1481
             | |||||||||| ||| || || || || ||||| ||| | || || ||||||||||||
Db      1993 GTTAATGCTTGCTAGGCAGACTGGCCTGACTAGGAGTCAGATTTCGAATTGGTTCATAAA 2052

Qy      1482 TGCTCGAGTTCGATTATGGAAGCCAATGGTAGAAGAAATGTACTTGGAAGAAGTGAAGAA 1541
             ||| || || ||  | ||||| |||||| | ||||| ||||||  |||||  |     |
Db      2053 TGCCCGTGTCCGCCTTTGGAAACCAATGATTGAAGACATGTACAAAGAAGAGATTGGTGA 2112

Qy      1542 TCAAGAACAAAACAGTACTAATACTTCAGGAGATAACAAAAACAAAGAGACCAATATAAG 1601
             |  ||| ||| ||   |    | | || |   ||      ||  |   ||  || ||
Db      2113 TTTAGAGCAAGACTCCAACTCTTCCTCTGACAATGCACCAAGAAGTAAGGATAAAATGGC 2172

Qy      1602 TGCTCCAAATGAAGAGAAACATCCAATTATTACTAG 1637
               || || | ||  || |  |||  | |   | ||
Db      2173 ATCTTCAGAAGACAAGGAGGATCTCAAAAACTCCAG 2208
```

```
RESULT 14
AK070465
LOCUS       AK070465               3030 bp    mRNA    linear   PLN 24-JUL-2003
DEFINITION  Oryza sativa (japonica cultivar-group) cDNA clone:J023054P21, full
            insert sequence.
ACCESSION   AK070465
VERSION     AK070465.1  GI:32980489
KEYWORDS    FLI_CDNA; CAP trapper.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1
  AUTHORS   The Rice Full-Length cDNA Consortium, National Institute of
            Agrobiological Sciences Rice Full-Length cDNA Project Team:,
            Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,
            Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,
            Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,
            Ohtsuki,K., Shishiki,T., Foundation of Advancement of International
            Science Genome Sequencing & Analysis Group:, Otomo,Y., Murakami,K.,
            Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,
            Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,
            Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J.,
            Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,
            Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN:,
            Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,
            Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,
            Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y.,
            Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
            Yoshino,M. and Hayashizaki,Y.
  TITLE     Collection, mapping, and annotation of over 28,000 cDNA clones from
            japonica rice
  JOURNAL   Science 301 (5631), 376-379 (2003)
  MEDLINE   22752273
   PUBMED   12869764
REFERENCE   2  (bases 1 to 3030)
  AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,
            Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W.,
            Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T.,
            Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K.,
            Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I.,
            Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M.,
            Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M.,
            Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M.,
            Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A.,
            Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M.,
            Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K.,
            Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H.,
            Osato,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K.,
            Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K.,
            Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S.,
            Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y.,
            Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A.,
            Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W.,
            Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and
            Yoshimura,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of
            Agrobiological Sciences, Department of Molecular Genetics, Head of
            Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
            305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp,
            Tel:81-29-838-7007, Fax:81-29-838-7007)
COMMENT     This clone is one of the 28K full-length cDNA clones from japonica
            rice.
            URL : http://cdna01.dna.affrc.go.jp/cDNA/
            NIAS Rice Full-Length cDNA Project Team: Kikuchi,S., Satoh,K.,
            Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J.,
            Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T.,
            Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and
            Yamamoto,M.
            FAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y.,
            Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M.,
            Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J.,
            Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S.,
            Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S.,
            Yoshimura,A., Matsubara,K. and Murakami,K.
            Genome Exploration Research Group in Riken Genomic Sciences Center
            and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K.,
            Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T.,
            Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K.,
            Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y.,
            Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J.,
            Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
            Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N.,
            Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H.,
            Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F.,
            Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K.,
            Yasunishi,A. and Hayashizaki,Y.
FEATURES             Location/Qualifiers
     source          1. .3030
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /clone="J023054P21"
ORIGIN

  Query Match            10.8%;  Score 295.6;  DB 8;  Length 3030;
  Best Local Similarity  60.3%;  Pred. No. 1.9e-49;
  Matches  528;  Conservative    0;  Mismatches  339;  Indels    9;  Gaps    2;

Qy        765 CTCTAAGTATCTGAAAGCTGCACAAGAGCTTCTTGATGAAGTTGTTAATATTGTTGGAAA 824
              ||| | |||||| || || |||||||| || |||||||| || || | | ||      |
Db       1394 CTCAAGGTATCTCAAGGCAGCACAAGAACTGCTTGATGAGGTCGTGAGTGTTTGGAAGAG 1453

Qy        825 AAGCATCAAAGGAGATGATCAAAAGAAGGATAATTCAATGAATAAAGAATCAATGCCTTT 884
              |   |    ||   | | |  || |||||| | | |||   ||   |||    |
Db       1454 CATAAAACAAAAGGCTCAGAAAGAGAAGGTTGAATCAGGAAAAGCAGATGGCAAA----- 1508

Qy        885 GGCTAGTGATGTCAACACTAATAGTTCTGGTGGTGGTGAAAGTAGCAGCAGGCAGAAAAA 944
              | | |||||      | || |||   ||||    |   | |
Db       1509 -GAGACTGATGGAGGGCCAAAAAGTGAGGGTGTATCTTCTAATCCACAAGAATCTGGCGC 1567

Qy        945 TGAAGTTGCTGTTGAGCTTACAACTGCTCAAAGACAAGAACTTCAAATGAAAAAAGCCAA 1004
              | | |  ||   |||||| | |||||| | |  ||||| ||||| |  ||||  || ||
Db       1568 TAATGCAGCACCTGAGCTTTCCACTGCTGAGAAGCAAGAGCTTCAGAACAAAATGGCAAA 1627

Qy       1005 GCTTCTTGCCATGCTTGAAGAGGTGGAGCAAAGGTACAGACAGTACCATCACCAAATGCA 1064
              ||  | || ||| | || |||||||| |   |  ||||  || ||  | ||||||||||
Db       1628 ACTGATGGCAATGTTGGATGAGGTGGACCGGAAATACAAGCATTATTACCACCAAATGCA 1687

Qy       1065 AATAATTGTATTATCATTTGAGCAAGTAGCAGGAATTGGATCAGCCAAATCATACACTCA 1124
              ||   | || | ||| |||||    ||||| |||   |||| || ||  | |||||
Db       1688 AACCGTGGTTTCATCGTTTGATGTGGTAGCTGGACCAGGATCTGCGAAGCCTTACACAGC 1747

Qy       1125 ATTAGCTTTGCATGCAATTTCGAAGCAATTCAGATGCCTAAAGGATGCAATTGCTGAGCA 1184
              | | ||| | ||  |||| ||   ||| ||  | ||||| |||||||| ||   ||| ||
Db       1748 AGTTGCTCTTCAGACAATCTCACGGCACTTTCGGTGCCTGAAGGATGCTATCAATGATCA 1807

Qy       1185 AGTAAAGGCGACGAGCAAGAGTTTAGGTGAAGAGGAAGGCTTGGGAGGGAAAATCGAAGG 1244
              | || |  |  || || |  | || || || ||  ||     || ||    | |
Db       1808 GATCAATGTTATCAGGAAAAAGCTTGGCGAGGAAGAGAACTCATCTGGCAAGGAGGGAAA 1867

Qy       1245 CTCAA---GACTCAAATTTGTGGACCATCATCTAAGGCAACAACGCGCGCTGCAACAGAT 1301
              |  ||    | |||    |   | || || || ||||||||||||||| ||  | ||||||
Db       1868 ATTAACCCGCCTCCGGTACATTGATCAGCAGCTAAGGCAACAACGTGCTTTCCAACAGTA 1927

Qy       1302 AGGAATGATGCAACCAAATGCTTGGAGACCCCAAAGAGGTTTACCTGAAAGAGCTGTCTC 1361
              || ||||| | || ||| || |||||||| || || ||  | |||||||    ||||  |
Db       1928 CGGTATGATTCCACAAAACGCCTGGAGACCACAGAGGGGACTGCCTGAAAACTCTGTTAC 1987

Qy       1362 TGTCCTTCGTGCTTGGCTTTTCGAGCATTTTCTTCATCCTTACCCAAAGGATTCAGACAA 1421
```

```
Db       1988 AATTCTTCGCGCTTGGCTGTTTGAACACTTCCTCCATCCGTACCCAAAAGATTCCGAGAA 2047

Qy       1422 AATCATGCTTGCTAAGCAAACGGGGCTAACAAGGAGCCAGGTGTCTAACTGGTTCATAAA 1481

Db       2048 GTTAATGCTTGCTAGGCAGACTGGCCTGACTAGGAGTCAGATTTCGAATTGGTTCATAAA 2107

Qy       1482 TGCTCGAGTTCGATTATGGAAGCCAATGGTAGAAGAAATGTACTTGGAAGAAGTGAAGAA 1541

Db       2108 TGCCCGTGTCCGCCTTTGGAAACCAATGATTGAAGACATGTACAAAGAAGAGATTGGTGA 2167

Qy       1542 TCAAGAACAAAACAGTACTAATACTTCAGGAGATAACAAAAACAAAGAGACCAATATAAG 1601

Db       2168 TTTAGAGCAAGACTCCAACTCTTCCTCTGACAATGCACCAAGAAGTAAGGATAAAATGGC 2227

Qy       1602 TGCTCCAAATGAAGAGAAACATCCAATTATTACTAG 1637

Db       2228 ATCTTCAGAAGACAAGGAGGATCTCAAAAACTCCAG 2263


RESULT 15
AF406703
LOCUS       AF406703               2065 bp    mRNA    linear   PLN 14-JUL-2003
DEFINITION  Solanum tuberosum BEL1-related homeotic protein 30 (Bel30) mRNA,
            complete cds.
ACCESSION   AF406703
VERSION     AF406703.1  GI:22652126
KEYWORDS    .
SOURCE      Solanum tuberosum (potato)
  ORGANISM  Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum.
REFERENCE   1  (bases 1 to 2065)
  AUTHORS   Chen,H., Rosin,F.M., Prat,S. and Hannapel,D.J.
  TITLE     Interacting transcription factors from the three-amino acid loop
            extension superclass regulate tuber formation
  JOURNAL   Plant Physiol. 132 (3), 1391-1404 (2003)
   PUBMED   12857821
REFERENCE   2  (bases 1 to 2065)
  AUTHORS   Chen,H., Rosin,F.M. and Hannapel,D.J.
  TITLE     A KNOX protein of potato interacts with several members of the TALE
            family of transcription factors
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 2065)
  AUTHORS   Chen,H., Rosin,F.M. and Hannapel,D.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-AUG-2001) Department of Horticulture, Iowa State
            University, 257 Horticulture Hall, Ames, IA 50011-1100, USA
FEATURES             Location/Qualifiers
     source          1. .2065
                     /organism="Solanum tuberosum"
                     /mol_type="mRNA"
                     /cultivar="Desiree"
                     /db_xref="taxon:4113"
     gene            1. .2065
                     /gene="Bel30"
     CDS             39. .1976
                     /gene="Bel30"
                     /note="StBel30"
                     /codon_start=1
                     /product="BEL1-related homeotic protein 30"
                     /protein_id="AAN03627.1"
                     /db_xref="GI:22652127"
                     /translation="MATYFPSPNNQRDADQTFQYFRQSLPESYSEASNAPENMMVFMN
                     YSSSGAYSDMLTGTSQQQHNCIDIPSIGATPFNTSQQEILSNLGGSQMGIQDFSSWRD
                     SRNEMLADNVFQVAQNVQGQGLSLSLGSNIPSGIGISHVQSQNPNQGGGFNMSFGDGD
                     NSQPKEQRNADYFPPDNPGRDLDAMKGYNSPYGTSSIARTIPSSKYLKAAQYLLDEVV
                     SVRKAIKEQNSKKELTKDSRESDVDSKNISSDTPANGGSNPHESKNNQSELSPTEKQE
                     VQNKLAKLLSMLDEIDRRYRQYYHQMQIVVSSFDVVAGEGAAKPYTALALQTISRHFR
                     CLRDAICDQIRASRRSLGEQDASENSKAIGISRLRFVDHHIRQQRALQQLGMMQQHAW
                     RPQRGLPESSVSVLRAWLFEHFLHPYPKDSDKIMLARQTGLTRSQVSNWFINARVRLW
                     KPMVEEMYKEEAGDAKIDSNSSSDVAPRLATKDSKVEERGELHQNAASEFEQYNSGQI
                     LESKSNHEADVEMEGASNAETQSQSGMENQTGEPLPAMDNCTLFQDAFVQSNDRFSEF
                     GSFGSGNVLPNGVSLTLGLQQGEGSNLPMSIETHVSYVPLRADDMYSTAPTTMVPETA
                     EFNCLDSGNRQQPFWLLPSAT"
ORIGIN

  Query Match           10.8%;  Score 295.2;  DB 8;  Length 2065;
  Best Local Similarity  61.0%;  Pred. No. 2.3e-49;
  Matches  499;  Conservative    0;  Mismatches  313;  Indels    6;  Gaps    1;

Qy        764 GCTCTAAGTATCTGAAAGCTGCACAAGAGCTTCTTGATGAAGTTGTTAATATTGTTGGAA 823

Db        646 GCTCGAAGTATTTGAAAGCAGCTCAATATTTGCTTGATGAGGTTGTTAGTGTCAGAAAGG 705

Qy        824 AAAGCATCAAAGGAGATGATCAAAAGAAGGATAATTCAATGAATAAAGAATCAATGCCTT 883

Db        706 CCATCAAGGAGCAAAATTCTAAGAAAGAGTTGACAAAGGATTCCAGAGAGTCTGATGTGG 765

Qy        884 TGGCTAGTGATGTCAACACTAATAGTTCTGGTGGTGGTGAAAGTAGCAGCAGGCAGAAAA 943

Db        766 ACTCGAAAAATATATCATCAGATACTCCTGCAAATGGGGGTTCAAATCCTCATGAGTCCA 825

Qy        944 ATGAAGTTGCTGTTGAGCTTACAACTGCTCAAAGACAAGAACTTCAAATGAAAAAAGCCA 1003

Db        826 AAAACAACCAAAGTGAACTTTCACCTACCGAGAAGCAAGAAGTGCAGAACAAACTGGCCA 885

Qy       1004 AGCTTCTTGCCATGCTTGAAGAGGTGGAGCAAAGGTACAGACAGTACCATCACCAAATGC 1063

Db        886 AACTTCTGTCAATGCTGGATGAGATTGATAGAAGGTACAGACAATATTATCATCAGATGC 945

Qy       1064 AAATAATTGTATTATCATTTGAGCAAGTAGCAGGAATTGGATCAGCCAAATCATACACTC 1123

Db        946 AAATAGTGGTTTCATCATTTGATGTGGTAGCTGGAGAAGGAGCAGCTAAACCATACACAG 1005

Qy       1124 AATTAGCTTTGCATGCAATTTCGAAGCAATTCAGATGCCTAAAGGATGCAATTGCTGAGC 1183

Db       1006 CTCTTGCTCTCCAGACAATTTCCCGACACTTCCGTTGCTTGCGTGATGCAATCTGCGATC 1065

Qy       1184 AAGTAAAGGCGACGAGCAAGAGTTTAGGTGAAGAGGAAGGCTTGGGAGGGA------AAA 1237

Db       1066 AGATTCGAGCATCACGAAGAAGTCTTGGAGAGCAAGATGCTTCAGAAAACAGCAAAGCGA 1125

Qy       1238 TCGAAGGCTCAAGACTCAAATTTGTGGACCATCATCTAAGGCAACAACGCGCGCTGCAAC 1297

Db       1126 TTGGAATATCACGCCTGCGTTTTGTGGATCATCATATTAGACAGCAGAGAGCCCTGCAGC 1185

Qy       1298 AGATAGGAATGATGCAACCAAATGCTTGGAGACCCCAAAGAGGTTTACCTGAAAGAGCTG 1357

Db       1186 AGCTTGGTATGATGCAACAACATGCCTGGAGGCCTCAGAGGGGATTGCCTGAAAGCTCTG 1245

Qy       1358 TCTCTGTCCTTCGTGCTTGGCTTTTCGAGCATTTTCTTCATCCTTACCCAAAGGATTCAG 1417

Db       1246 TTTCAGTTTTGCGTGCTTGGCTCTTTGAGCACTTTCTTCATCCCTACCCGAAAGATTCTG 1305

Qy       1418 ACAAAATCATGCTTGCTAAGCAAACGGGGCTAACAAGGAGCCAGGTGTCTAACTGGTTCA 1477

Db       1306 ACAAAATTATGCTAGCAAGGCAAACTGGCTTAACGAGAAGTCAGGTATCAAATTGGTTCA 1365

Qy       1478 TAAATGCTCGAGTTCGATTATGGAAGCCAATGGTAGAAGAAATGTACTTGGAAGAAGTGA 1537

Db       1366 TAAATGCACGGGTGCGTCTTTGGAAACCCATGGTTGAGGAAATGTACAAAGAAGAGGCTG 1425

Qy       1538 AGAATCAAGAACAAAACAGTACTAATACTTCAGGAGAT 1575

Db       1426 GTGATGCTAAAATAGACTCAAATTCTTCATCGGATGTT 1463


Search completed: July 29, 2005, 11:50:00
Job time : 11605 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 29, 2005, 05:17:43 ; Search time 1407 Seconds
(without alignments)
11507.102 Million cell updates/sec

Title: US-10-624-201A-1
Perfect score: 2735
Sequence: 1 catgcagagataaaaatata.....gcaaaaaaaaaaaaaaaaa 2735

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	324.6	11.9	2043	6	ABZ13532	Abz13532 Arabidops
2	324.6	11.9	2043	9	ADB23149	Adb23149 Environme
3	324.6	11.9	2385	10	ADD30303	Add30303 Plant yie
4	324.6	11.9	2385	12	ADI44208	Adi44208 Plant tra
5	278.8	10.2	1929	12	ADQ37104	Adq37104 Cell prol
6	261	9.5	2031	3	AAC36460	Aac36460 Arabidops
7	259.4	9.5	1449	3	AAC43406	Aac43406 Arabidops
8	259.4	9.5	1449	6	ADG88236	Adg88236 A. thalia
9	244.4	8.9	1971	12	ADO62706	Ado62706 Transcrip
10	237.6	8.7	706	6	ABK82114	Abk82114 DNA encod
11	234.6	8.6	2025	3	AAC39138	Aac39138 Arabidops
12	234.4	8.6	1575	4	AAD06493	Aad06493 Arabidops
13	234.4	8.6	1575	12	ADO61968	Ado61968 Transcrip
14	234.4	8.6	1824	3	AAC36745	Aac36745 Arabidops
15	231.8	8.5	2458	12	ADO62707	Ado62707 Transcrip
16	230.2	8.4	1617	12	ADO63042	Ado63042 Transcrip
17	225.4	8.2	1599	6	ABZ14349	Abz14349 Arabidops
18	224	8.2	1884	6	ABZ13100	Abz13100 Arabidops
19	224	8.2	1983	4	AAD06494	Aad06494 Arabidops
20	224	8.2	1983	10	ADC46618	Adc46618 Thalecres
21	224	8.2	1983	10	ADD30786	Add30786 Plant yie
22	224	8.2	1983	12	ADI44018	Adi44018 Plant tra
23	224	8.2	1983	12	ADO02224	Ado02224 Thalecres
24	222.6	8.1	2545	4	AAD06479	Aad06479 Arabidops
25	222.6	8.1	2545	10	ADE37296	Ade37296 Plant yie
26	222.6	8.1	2545	12	ADI43866	Adi43866 Plant tra
27	222.6	8.1	2545	12	ADO01854	Ado01854 Thalecres
28	213.4	7.8	1485	3	AAC50215	Aac50215 Arabidops
29	207	7.6	1296	12	ADO61794	Ado61794 Transcrip
30	201.4	7.4	626	5	AAH87740	Aah87740 Peppermin
31	187.2	6.8	1836	6	ABZ12846	Abz12846 Arabidops
32	187.2	6.8	1836	6	ADG88038	Adg88038 A. thalia
33	187.2	6.8	2405	4	AAD06449	Aad06449 Arabidops
34	187.2	6.8	2405	10	ADE37116	Ade37116 Plant yie
35	187.2	6.8	2405	12	ADI43868	Adi43868 Plant tra
36	187.2	6.8	2405	12	ADO01856	Ado01856 Thalecres
37	183.8	6.7	529	3	AAC55946	Aac55946 Eucalyptu
38	182.8	6.7	662	13	ADR65754	Adr65754 Cotton cD
39	180.4	6.6	592	13	ACN59439	Acn59439 Cotton gy
40	176.8	6.5	2034	3	AAC50226	Aac50226 Arabidops
41	170.2	6.2	276	12	ADQ05068	Adq05068 Soybean h
42	166	6.1	894	12	ADQ37116	Adq37116 Cell prol
43	164.6	6.0	343	3	AAC56856	Aac56856 Pinus rad
44	159.8	5.8	323	3	AAC56287	Aac56287 Pinus rad
45	158	5.8	833	3	AAC34112	Aac34112 Arabidops

ALIGNMENTS

RESULT 1
ABZ13532
ID ABZ13532 standard; DNA; 2043 BP.
XX
AC ABZ13532;
XX
DT 21-JAN-2003 (first entry)
XX
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 1337.
XX
KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX
OS Arabidopsis thaliana.
XX
PN WO200216655-A2.
XX
PD 28-FEB-2002.
XX
PF 24-AUG-2001; 2001WO-US026685.
XX
PR 24-AUG-2000; 2000US-0227866P.
PR 26-JAN-2001; 2001US-0264647P.
PR 22-JUN-2001; 2001US-0300111P.
XX
PA (SCRI) SCRIPPS RES INST.
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Harper JF, Kreps J, Wang X, Zhu T;
XX
DR WPI; 2002-304127/34.
XX
PT Identifying a stress condition to which a plant cell has been exposed and
PT producing plants with increased tolerance to these abiotic stresses.
XX
PS Claim 144; SEQ ID NO 1337; 577pp + Sequence Listing; English.
XX
CC The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising: (a) contacting nucleic acid
CC representative of expressed polynucleotides in the plant cell with an
CC array or probes representative of the plant cell genome; and (b)
CC detecting a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants

```
CC   with increased tolerance to abiotic stress. The present sequence is that
CC   of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
CC   in methods of the invention. Note: The sequence data for this patent is
CC   not represented in the printed specification but is based on sequence
CC   information supplied to Derwent by the European Patent Office
XX
SQ   Sequence 2043 BP; 627 A; 507 C; 507 G; 402 T; 0 U; 0 Other;

  Query Match             11.9%;  Score 324.6;  DB 6;  Length 2043;
  Best Local Similarity   65.2%;  Pred. No. 3.8e-64;
  Matches  534;  Conservative    0;  Mismatches  264;  Indels   21;  Gaps    3;

Qy        758 TTTTAGGCTCTAAGTATCTGAAAGCTGCACAAGAGCTTCTTGATGAAGTTGTTAATATTG 817
              || |  |||| |||||  |||| || ||||||||||||||||| |||||  || ||   ||
Db        569 TTGTTAGCTCCAAGTACTTGAAGGCAGCACAAGAGCTTCTTGACGAAGTAGTCAACGCTG 628

Qy        818 TTGGAAAAAGCATCAAAGGAGATGATCAAAAGAAGGATAATTCAATGAATAAAGAATCAA 877
                |    |   ||| || |   |    |||              || |  ||       ||
Db        629 ATTCCGATGACATGAACGCTAAATCCCAACTATTCTCATCGAAAAAGGGTAGTTGCGGAA 688

Qy        878 TGCCTTTGGCTAGTGATGTCAACACTAATAGTTCTGGTGGTGGTGAAAGTAGCAGCAGGC 937
                  ||   |  |       |     |  |||| || |  |      |  || |
Db        689 ATGATAAACCTGTCGGAGAATCATCGGCCGGCGCTGGAGGAGAAGGTTCCGGTGGCGGAG 748

Qy        938 AGAAAAATG------AAGTTGCTGTTGAGCTTACAACTGCTCAAAGACAAGAACTTCAAA 991
                 ||     |       ||  | | || |||||     || ||  | |||||||| | || |
Db        749 CAGAAGCAGCCGGGAAACGTCCGGTGGAGCTAGGCACGGCAGAGAGACAAGAAATACAGA 808

Qy        992 TGAAAAAAGCCAAGCTTCTTGCCATGCTTGAAGAGGTGGAGCAAAGGTACAGACAGTACC 1051
              |||| ||||| || |||  |  ||||||| | ||||||||||| || || ||||||||||
Db        809 TGAAGAAAGCAAAACTTAGTAACATGCTTCATGAGGTGGAGCAGAGATATAGACAGTACC 868

Qy       1052 ATCACCAAATGCAAATAATTGTATTATCATTTGAGCAAGTAGCAGGAATTGGATCAGCCA 1111
              | || || ||||| ||  |  | |  || || |||||||  ||||| || |||||||| |
Db        869 ACCAGCAGATGCAGATGGTGATCTCTTCGTTCGAGCAAGCGGCAGGGATAGGATCAGCGA 928

Qy       1112 AATCATACACTCAATTAGCTTTGCATGCAATTTCGAAGCAATTCAGATGCCTAAAGGATG 1171
              | ||||||||      |||| ||| |  | || || |  || ||| | ||| | || || | |
Db        929 AGTCATACACGTCGCTAGCATTGAAAACCATATCAAGACAGTTCCGTTGCTTGAAAGAGG 988

Qy       1172 CAATTGCTGAGCAAGTAAAGGCGACGAGCAAGAGTTTAGGTGAAGAGGAAGGCTTG---- 1227
              | || ||||  ||  |||| ||| | | ||||||| | || || || ||      ||
Db        989 CGATCGCTGGTCAGATAAAAGCGGCCAACAAGAGTCTTGGGGAGGAAGATTCAGTGTCTG 1048

Qy       1228 --GGAGGGAAAATCGAAGGCTCAAGACTCAAATTTGTGGACCATCATCTAAGGCAACAAC 1285
                |  ||||   | || || || || ||||| || |||||||  ||  | || || |||
Db       1049 GTGTTGGGAGGTTTGAGGGGTCGAGGCTCAAGTTCGTGGACCACCACTTGAGACAGCAAA 1108

Qy       1286 GCGCGCTGCAACAGATAGGAATGATGCAAC---------CAAATGCTTGGAGACCCCAAA 1336
              | || || |||||  | |||||||| ||||          ||||||||||||| |||
Db       1109 GAGCTCTTCAACAACTGGGAATGATTCAACATCCTTCCAATAATGCTTGGAGACCTCAAC 1168

Qy       1337 GAGGTTTACCTGAAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTTTTCGAGCATTTTCTTC 1396
              | ||| | || ||| |||| ||||| || || |||||||||||  ||||| || |||||||
Db       1169 GTGGTCTCCCAGAACGAGCCGTCTCAGTTCTCCGTGCTTGGCTCTTCGAACACTTTCTTC 1228

Qy       1397 ATCCTTACCCAAAGGATTCAGACAAAATCATGCTTGCTAAGCAAACGGGGCTAACAAGGA 1456
              |||| ||||| |||||||| |||||   |||||| |||||||||| || || ||  | |
Db       1229 ATCCATACCCTAAGGATTCGGACAAGCACATGCTAGCTAAGCAAACAGGACTCACTCGTA 1288

Qy       1457 GCCAGGTGTCTAACTGGTTCATAAATGCTCGAGTTCGATTATGGAAGCCAATGGTAGAAG 1516
              |||||||||| |||||||| ||||| ||  |||||||  |||||||| |||||||| || |
Db       1289 GCCAGGTGTCGAACTGGTTTATAAACGCGAGAGTTCGGTTATGGAAACCAATGGTGGAGG 1348

Qy       1517 AAATGTACTTGGAAGAAGTGAAGAATCAAGAACAAAACA 1555
              | |||||| |||| ||| ||||| | || | | | ||||
Db       1349 AGATGTACATGGAGGAAATGAAGGAGCAGGCAAAGAACA 1387


RESULT 2
ADB23149
ID   ADB23149 standard; DNA; 2043 BP.
XX
AC   ADB23149;
XX
DT   20-NOV-2003  (first entry)
XX
DE   Environmental stress-responsive promoter-related DNA, SEQ ID 153.
XX
KW   Plant; environmental stress; promoter; gene; ds.
XX
OS   Arabidopsis thaliana.
XX
PN   WO2003044190-A1.
XX
PD   30-MAY-2003.
XX
PF   15-NOV-2002; 2002WO-JP011955.
XX
PR   19-NOV-2001; 2001JP-00353038.
PR   29-JAN-2002; 2002JP-00020329.
XX
PA   (RIKE ) RIKEN KK.
XX
PI   Shinozaki K,  Seki M,  Fujita M;
XX
DR   WPI; 2003-468771/44.
DR   P-PSDB; ADB23150.
XX
PT   Environmental stress-responsive promoters, useful in constructing and
PT   breeding stress-resistant plants for applications in agriculture.
XX
PS   Claim 10; Page 470-475; 520pp; Japanese.
XX
CC   The present invention relates to novel environmental stress-responsive
CC   promoters (ADB23163-ADB23252) from Arabidopsis thaliana. The promoters
CC   are useful in constructing and breeding stress-resistant plants for
CC   applications in agriculture and horticulture. The present sequence was
CC   used to illustrate the invention.
XX
SQ   Sequence 2043 BP; 627 A; 507 C; 507 G; 402 T; 0 U; 0 Other;

  Query Match             11.9%;  Score 324.6;  DB 9;  Length 2043;
  Best Local Similarity   65.2%;  Pred. No. 3.8e-64;
  Matches  534;  Conservative    0;  Mismatches  264;  Indels   21;  Gaps    3;

Qy        758 TTTTAGGCTCTAAGTATCTGAAAGCTGCACAAGAGCTTCTTGATGAAGTTGTTAATATTG 817
              || |  |||| |||||  |||| || ||||||||||||||||| |||||  || ||   ||
Db        569 TTGTTAGCTCCAAGTACTTGAAGGCAGCACAAGAGCTTCTTGACGAAGTAGTCAACGCTG 628

Qy        818 TTGGAAAAAGCATCAAAGGAGATGATCAAAAGAAGGATAATTCAATGAATAAAGAATCAA 877
                |    |   ||| || |   |    |||              || |  ||       ||
Db        629 ATTCCGATGACATGAACGCTAAATCCCAACTATTCTCATCGAAAAAGGGTAGTTGCGGAA 688

Qy        878 TGCCTTTGGCTAGTGATGTCAACACTAATAGTTCTGGTGGTGGTGAAAGTAGCAGCAGGC 937
                  ||   |  |       |     |  |||| || |  |      |  || |
Db        689 ATGATAAACCTGTCGGAGAATCATCGGCCGGCGCTGGAGGAGAAGGTTCCGGTGGCGGAG 748

Qy        938 AGAAAAATG------AAGTTGCTGTTGAGCTTACAACTGCTCAAAGACAAGAACTTCAAA 991
                 ||     |       ||  | | || |||||     || ||  | |||||||| | || |
Db        749 CAGAAGCAGCCGGGAAACGTCCGGTGGAGCTAGGCACGGCAGAGAGACAAGAAATACAGA 808

Qy        992 TGAAAAAAGCCAAGCTTCTTGCCATGCTTGAAGAGGTGGAGCAAAGGTACAGACAGTACC 1051
              |||| ||||| || |||  |  ||||||| | ||||||||||| || || ||||||||||
Db        809 TGAAGAAAGCAAAACTTAGTAACATGCTTCATGAGGTGGAGCAGAGATATAGACAGTACC 868

Qy       1052 ATCACCAAATGCAAATAATTGTATTATCATTTGAGCAAGTAGCAGGAATTGGATCAGCCA 1111
              | || || ||||| ||  |  | |  || || |||||||  ||||| || |||||||| |
Db        869 ACCAGCAGATGCAGATGGTGATCTCTTCGTTCGAGCAAGCGGCAGGGATAGGATCAGCGA 928

Qy       1112 AATCATACACTCAATTAGCTTTGCATGCAATTTCGAAGCAATTCAGATGCCTAAAGGATG 1171
              | ||||||||      |||| ||| |  | || || |  || ||| | ||| | || || | |
Db        929 AGTCATACACGTCGCTAGCATTGAAAACCATATCAAGACAGTTCCGTTGCTTGAAAGAGG 988

Qy       1172 CAATTGCTGAGCAAGTAAAGGCGACGAGCAAGAGTTTAGGTGAAGAGGAAGGCTTG---- 1227
              | || ||||  ||  |||| ||| | | ||||||| | || || || ||      ||
```

```
Db        989 CGATCGCTGGTCAGATAAAAGCGGCCAACAAGAGTCTTGGGGAGGAAGATTCAGTGTCTG 1048

Qy       1228 --GGAGGGAAAATCGAAGGCTCAAGACTCAAATTTGTGGACCATCATCTAAGGCAACAAC 1285
                |  ||||   | || || || || ||||| || |||||||| ||  | || || |||
Db       1049 GTGTTGGGAGGTTTGAGGGGTCGAGGCTCAAGTTCGTGGACCACCACTTGAGACAGCAAA 1108

Qy       1286 GCGCGCTGCAACAGATAGGAATGATGCAAC---------CAAATGCTTGGAGACCCCAAA 1336
                | || || |||||  | |||||||| ||||           |||||||||||||| |||
Db       1109 GAGCTCTTCAACAACTGGGAATGATTCAACATCCTTCCAATAATGCTTGGAGACCTCAAC 1168

Qy       1337 GAGGTTTACCTGAAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTTTTCGAGCATTTTCTTC 1396
                | ||| | || ||| |||| ||||| || || |||||||||||| ||||| || |||||||
Db       1169 GTGGTCTCCCAGAACGAGCCGTCTCAGTTCTCCGTGCTTGGCTCTTCGAACACTTTCTTC 1228

Qy       1397 ATCCTTACCCAAAGGATTCAGACAAAATCATGCTTGCTAAGCAAACGGGGCTAACAAGGA 1456
                |||| ||||| |||||||| |||||    |||||| |||||||||| || || ||  | |
Db       1229 ATCCATACCCTAAGGATTCGGACAAGCACATGCTAGCTAAGCAAACAGGACTCACTCGTA 1288

Qy       1457 GCCAGGTGTCTAACTGGTTCATAAATGCTCGAGTTCGATTATGGAAGCCAATGGTAGAAG 1516
                |||||||||| |||||||| ||||| ||  ||||||| |||||||| |||||||| || |
Db       1289 GCCAGGTGTCGAACTGGTTTATAAACGCGAGAGTTCGGTTATGGAAACCAATGGTGGAGG 1348

Qy       1517 AAATGTACTTGGAAGAAGTGAAGAATCAAGAACAAAACA 1555
                | |||||| |||| ||| ||||| | || | | | ||||
Db       1349 AGATGTACATGGAGGAAATGAAGGAGCAGGCAAAGAACA 1387


RESULT 3
ADD30303
ID   ADD30303 standard; cDNA; 2385 BP.
XX
AC   ADD30303;
XX
DT   15-JAN-2004  (first entry)
XX
DE   Plant yield-related polynucleotide clone G1589.
XX
KW   ds; transcription factor; transgenic plant; growth rate; senescence;
KW   seed germination rate; plant vigor; seedling vigor.
XX
OS   Arabidopsis thaliana.
XX
PN   WO2003013227-A2.
XX
PD   20-FEB-2003.
XX
PF   09-AUG-2002; 2002WO-US025805.
XX
PR   09-AUG-2001; 2001US-0310847P.
PR   19-NOV-2001; 2001US-0336049P.
PR   11-DEC-2001; 2001US-0338692P.
PR   14-JUN-2002; 2002US-00171468.
XX
PA   (MEND-) MENDEL BIOTECHNOLOGY INC.
XX
PI   Ratcliffe O,  Riechmann JL,  Adam LJ,  Dubell AT,  Heard JE;
PI   Pilgrim ML,  Jiang C,  Reuber TL,  Creelman RA,  Pineda O,  Yu G;
PI   Broun PE;
XX
DR   WPI; 2003-248221/24.
DR   P-PSDB; ADD30304.
XX
PT   New plant transcription factor polynucleotides and polypeptides, useful
PT   in producing transgenic plants with commercially valuable properties,
PT   such as an alteration in a plant growth characteristic, e.g. growth rate
PT   or apomixis.
XX
PS   Disclosure; SEQ ID NO 332; 454pp; English.
XX
CC   The invention relates to a number of isolated Arabidopsis thaliana cDNA
CC   sequences and their encoded proteins which are especially transcription
CC   factor related cDNA's and proteins. The isolated or recombinant plant
CC   transcription factor polynucleotides and polypeptides are useful in
CC   producing transgenic plants with commercially valuable properties, i.e.
CC   modified or altered desirable traits as compared to a reference plant,
CC   such as an alteration in a plant growth characteristic, e.g. growth rate,
CC   germination rate of seeds, vigor of plants and seedlings, or leaf and
CC   flower senescence. Sequence information related to the polynucleotides
CC   and polypeptides can also be used in bioinformatic search methods. The
CC   transgenic plant is useful for growing a progeny plant from a parent
CC   plant. This sequence represents one of the cDNAs of the invention.
XX
SQ   Sequence 2385 BP; 748 A; 578 C; 565 G; 494 T; 0 U; 0 Other;

  Query Match           11.9%;  Score 324.6;  DB 10;  Length 2385;
  Best Local Similarity   65.2%;  Pred. No. 4e-64;
  Matches  534;  Conservative    0;  Mismatches  264;  Indels   21;  Gaps    3;

Qy        758 TTTTAGGCTCTAAGTATCTGAAAGCTGCACAAGAGCTTCTTGATGAAGTTGTTAATATTG 817
              || |  |||| |||||  |||| || |||||||||||||||| |||||  || ||    ||
Db        747 TTGTTAGCTCCAAGTACTTGAAGGCAGCACAAGAGCTTCTTGACGAAGTAGTCAACGCTG 806

Qy        818 TTGGAAAAAGCATCAAAGGAGATGATCAAAAGAAGGATAATTCAATGAATAAAGAATCAA 877
              |    |  ||| || |   |   |||                 || |  ||       ||
Db        807 ATTCCGATGACATGAACGCTAAATCCCAACTATTCTCATCGAAAAAGGGTAGTTGCGGAA 866

Qy        878 TGCCTTTGGCTAGTGATGTCAACACTAATAGTTCTGGTGGTGGTGAAAGTAGCAGCAGGC 937
                  |    ||   |  |       |    |||| || |  |       |  || |
Db        867 ATGATAAACCTGTCGGAGAATCATCGGCCGGCGCTGGAGGAGAAGGTTCCGGTGGCGGAG 926

Qy        938 AGAAAAATG------AAGTTGCTGTTGAGCTTACAACTGCTCAAAGACAAGAACTTCAAA 991
                 ||  |      ||  | | || |||||    || ||  | |||||||||  | || |
Db        927 CAGAAGCAGCCGGGAAACGTCCGGTGGAGCTAGGCACGGCAGAGAGACAAGAAATACAGA 986

Qy        992 TGAAAAAAGCCAAGCTTCTTGCCATGCTTGAAGAGGTGGAGCAAAGGTACAGACAGTACC 1051
              |||| ||||| || |||  |  ||||||| | ||||||||||| || || ||||||||||
Db        987 TGAAGAAAGCAAAACTTAGTAACATGCTTCATGAGGTGGAGCAGAGATATAGACAGTACC 1046

Qy       1052 ATCACCAAATGCAAATAATTGTATTATCATTTGAGCAAGTAGCAGGAATTGGATCAGCCA 1111
              | || || ||||| ||  |  | |  || || |||||||  ||||| || |||||||| |
Db       1047 ACCAGCAGATGCAGATGGTGATCTCTTCGTTCGAGCAAGCGGCAGGGATAGGATCAGCGA 1106

Qy       1112 AATCATACACTCAATTAGCTTTGCATGCAATTTCGAAGCAATTCAGATGCCTAAAGGATG 1171
              | ||||||||      |||| ||| |  | || || |  || ||| | ||| | || || |
Db       1107 AGTCATACACGTCGCTAGCATTGAAAACCATATCAAGACAGTTCCGTTGCTTGAAAGAGG 1166

Qy       1172 CAATTGCTGAGCAAGTAAAGGCGACGAGCAAGAGTTTAGGTGAAGAGGAAGGCTTG---- 1227
              | || |||| ||  |||| ||| | | ||||||| | || || || ||      ||
Db       1167 CGATCGCTGGTCAGATAAAAGCGGCCAACAAGAGTCTTGGGGAGGAAGATTCAGTGTCTG 1226

Qy       1228 --GGAGGGAAAATCGAAGGCTCAAGACTCAAATTTGTGGACCATCATCTAAGGCAACAAC 1285
                |  ||||   | || || || || ||||| || |||||||| ||  | ||| || |||
Db       1227 GTGTTGGGAGGTTTGAGGGGTCGAGGCTCAAGTTCGTGGACCACCACTTGAGACAGCAAA 1286

Qy       1286 GCGCGCTGCAACAGATAGGAATGATGCAAC---------CAAATGCTTGGAGACCCCAAA 1336
              | || || |||||  | |||||||| ||||           |||||||||||||| |||
Db       1287 GAGCTCTTCAACAACTGGGAATGATTCAACATCCTTCCAATAATGCTTGGAGACCTCAAC 1346

Qy       1337 GAGGTTTACCTGAAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTTTTCGAGCATTTTCTTC 1396
              | ||| | || ||| |||| ||||| || || |||||||||||| ||||| || |||||||
Db       1347 GTGGTCTCCCAGAACGAGCCGTCTCAGTTCTCCGTGCTTGGCTCTTCGAACACTTTCTTC 1406

Qy       1397 ATCCTTACCCAAAGGATTCAGACAAAATCATGCTTGCTAAGCAAACGGGGCTAACAAGGA 1456
              |||| ||||| |||||||| |||||    |||||| |||||||||||| || ||  | |
Db       1407 ATCCATACCCTAAGGATTCGGACAAGCACATGCTAGCTAAGCAAACAGGACTCACTCGTA 1466

Qy       1457 GCCAGGTGTCTAACTGGTTCATAAATGCTCGAGTTCGATTATGGAAGCCAATGGTAGAAG 1516
              |||||||||| |||||||| ||||| ||  ||||||| |||||||| |||||||| || |
Db       1467 GCCAGGTGTCGAACTGGTTTATAAACGCGAGAGTTCGGTTATGGAAACCAATGGTGGAGG 1526

Qy       1517 AAATGTACTTGGAAGAAGTGAAGAATCAAGAACAAAACA 1555
              | |||||| |||| ||| ||||| | || | | | ||||
Db       1527 AGATGTACATGGAGGAAATGAAGGAGCAGGCAAAGAACA 1565


RESULT 4
```

```
ADI44208
ID   ADI44208 standard; DNA; 2385 BP.
XX
AC   ADI44208;
XX
DT   22-APR-2004  (first entry)
XX
DE   Plant transcription factor related polynucleotide #1693.
XX
KW   transgenic; plant; enhanced tolerance to abiotic stress;
KW   glyphosphate tolerance; hormone sensitivity; disease resistance;
KW   sugar sensing; flowering; flower structure; stem bifurcation;
KW   branching pattern; apical dominance; trichome; stem morphology;
KW   root growth; root hair; seed development; cell proliferation;
KW   cell differentiation; premature senescence; necrosis; plant size;
KW   leaf morphology; seed morphology; seed biochemistry; root anthocyanin;
KW   plant anthocyanin; light response; shade avoidance; bioinformatic;
KW   transcription factor; ds.
XX
OS   Unidentified.
XX
PN   US2004019927-A1.
XX
PD   29-JAN-2004.
XX
PF   25-FEB-2003; 2003US-00374780.
XX
PR   18-APR-2001; 2001US-00837944.
XX
PA   (SHER/) SHERMAN B K.
PA   (RIEC/) RIECHMANN J L.
PA   (JIAN/) JIANG C.
PA   (HEAR/) HEARD J E.
PA   (HAAK/) HAAKE V.
PA   (CREE/) CREELMAN R A.
PA   (RATC/) RATCLIFFE O.
PA   (ADAM/) ADAM L J.
PA   (REUB/) REUBER T L.
PA   (KEDD/) KEDDIE J.
PA   (BROU/) BROUN P E.
PA   (PILG/) PILGRIM M L.
PA   (DUBE/) DUBELL A N.
PA   (PINE/) PINEDA O.
PA   (YUGG/) YU G.
XX
PI   Sherman BK,  Riechmann JL,  Jiang C,  Heard JE,  Haake V;
PI   Creelman RA,  Ratcliffe O,  Adam LJ,  Reuber TL,  Keddie J,  Broun PE;
PI   Pilgrim ML,  Dubell AN,  Pineda O,  Yu G;
XX
DR   WPI; 2004-132245/13.
DR   P-PSDB; ADI44209.
XX
PT   New transgenic plant comprising a recombinant polynucleotide of any one
PT   of more than 500 nucleotide sequences, useful in bioinformatic search
PT   methods.
XX
PS   Disclosure; SEQ ID NO 2671; 435pp; English.
XX
CC   The invention describes a transgenic plant comprising a recombinant
CC   polynucleotide of any one of more than 500 nucleotide sequences fully
CC   defined in the specification or its complement. The method of the
CC   invention can be used to produced a plant having altered traits such as:
CC   enhanced tolerance to abiotic stress; glyphosphate tolerance; hormone
CC   sensitivity; disease resistance; sugar sensing; early or late flowering;
CC   altered flower structure, change in stem bifurcations, altered branching
CC   pattern, reduced apical dominance, reduced trichome density; lack of
CC   trichomes; reduced ectopic trichome development; altered trichome
CC   development; increase in trichome number; altered stem morphology;
CC   increased root growth; increased root hairs; altered seed development;
CC   altered cell proliferation or cell differentiation; rapid development;
CC   premature senescence; increased necrosis; increase in seedling or plant
CC   size; decreased plant size; leaf morphology; seed morphology; seed
CC   biochemistry; increase in root anthocyanins; increase in plant
CC   anthocyanins, or alteration in light response or shade avoidance. The
CC   transgenic plant, polynucleotides and polypeptides are useful in
CC   bioinformatic search methods. This sequence represents a plant
CC   transcription factor related polynucleotide.
XX
SQ   Sequence 2385 BP; 748 A; 578 C; 565 G; 494 T; 0 U; 0 Other;

  Query Match            11.9%;  Score 324.6;  DB 12;  Length 2385;
  Best Local Similarity  65.2%;  Pred. No. 4e-64;
  Matches  534;  Conservative    0;  Mismatches  264;  Indels   21;  Gaps    3;

Qy       758 TTTTAGGCTCTAAGTATCTGAAAGCTGCACAAGAGCTTCTTGATGAAGTTGTTAATATTG 817
Db       747 TTGTTAGCTCCAAGTACTTGAAGGCAGCACAAGAGCTTCTTGACGAAGTAGTCAACGCTG 806

Qy       818 TTGGAAAAAGCATCAAAGGAGATGATCAAAAGAAGGATAATTCAATGAATAAAGAATCAA 877
Db       807 ATTCCGATGACATGAACGCTAAATCCCAACTATTCTCATCGAAAAAGGGTAGTTGCGGAA 866

Qy       878 TGCCTTTGGCTAGTGATGTCAACACTAATAGTTCTGGTGGTGGTGAAAGTAGCAGCAGGC 937
Db       867 ATGATAAACCTGTCGGAGAATCATCGGCCGGCGCTGGAGGAGAAGGTTCCGGTGGCGGAG 926

Qy       938 AGAAAAATG------AAGTTGCTGTTGAGCTTACAACTGCTCAAAGACAAGAACTTCAAA 991
Db       927 CAGAAGCAGCCGGGAAACGTCCGGTGGAGCTAGGCACGGCAGAGAGACAAGAAATACAGA 986

Qy       992 TGAAAAAAGCCAAGCTTCTTGCCATGCTTGAAGAGGTGGAGCAAAGGTACAGACAGTACC 1051
Db       987 TGAAGAAAGCAAAACTTAGTAACATGCTTCATGAGGTGGAGCAGAGATATAGACAGTACC 1046

Qy      1052 ATCACCAAATGCAAATAATTGTATTATCATTTGAGCAAGTAGCAGGAATTGGATCAGCCA 1111
Db      1047 ACCAGCAGATGCAGATGGTGATCTCTTCGTTCGAGCAAGCGGCAGGGATAGGATCAGCGA 1106

Qy      1112 AATCATACACTCAATTAGCTTTGCATGCAATTTCGAAGCAATTCAGATGCCTAAAGGATG 1171
Db      1107 AGTCATACACGTCGCTAGCATTGAAAACCATATCAAGACAGTTCCGTTGCTTGAAAGAGG 1166

Qy      1172 CAATTGCTGAGCAAGTAAAGGCGACGAGCAAGAGTTTAGGTGAAGAGGAAGGCTTG---- 1227
Db      1167 CGATCGCTGGTCAGATAAAAGCGGCCAACAAGAGTCTTGGGGAGGAAGATTCAGTGTCTG 1226

Qy      1228 --GGAGGGAAAATCGAAGGCTCAAGACTCAAATTTGTGGACCATCATCTAAGGCAACAAC 1285
Db      1227 GTGTTGGGAGGTTTGAGGGGTCGAGGCTCAAGTTCGTGGACCACCACTTGAGACAGCAAA 1286

Qy      1286 GCGCGCTGCAACAGATAGGAATGATGCAAC---------CAAATGCTTGGAGACCCCAAA 1336
Db      1287 GAGCTCTTCAACAACTGGGAATGATTCAACATCCTTCCAATAATGCTTGGAGACCTCAAC 1346

Qy      1337 GAGGTTTACCTGAAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTTTTCGAGCATTTTCTTC 1396
Db      1347 GTGGTCTCCCAGAACGAGCCGTCTCAGTTCTCCGTGCTTGGCTCTTCGAACACTTTCTTC 1406

Qy      1397 ATCCTTACCCAAAGGATTCAGACAAAATCATGCTTGCTAAGCAAACGGGGCTAACAAGGA 1456
Db      1407 ATCCATACCCTAAGGATTCGGACAAGCACATGCTAGCTAAGCAAACAGGACTCACTCGTA 1466

Qy      1457 GCCAGGTGTCTAACTGGTTCATAAATGCTCGAGTTCGATTATGGAAGCCAATGGTAGAAG 1516
Db      1467 GCCAGGTGTCGAACTGGTTTATAAACGCGAGAGTTCGGTTATGGAAACCAATGGTGGAGG 1526

Qy      1517 AAATGTACTTGGAAGAAGTGAAGAATCAAGAACAAAACA 1555
Db      1527 AGATGTACATGGAGGAAATGAAGGAGCAGGCAAAGAACA 1565


RESULT 5
ADQ37104
ID   ADQ37104 standard; DNA; 1929 BP.
XX
AC   ADQ37104;
```

XX
DT 07-OCT-2004 (first entry)
XX
DE Cell proliferation-related nucleic acid sequence #132.
XX
KW cell proliferation related polypeptide; cell proliferation; senescence;
KW differentiation; stress response; ds.
XX
OS Oryza sativa.
XX
PN WO2004061122-A2.
XX
PD 22-JUL-2004.
XX
PF 23-DEC-2003; 2003WO-US041200.
XX
PR 26-DEC-2002; 2002US-0436565P.
XX
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Cooper B;
XX
DR WPI; 2004-534388/51.
XX
PT New nucleic acid molecule encoding a cell proliferation-related
PT polypeptide, useful for modulating cell proliferation, senescence,
PT differentiation, development, and stress response in plants, and for
PT producing enhanced food crops.
XX
PS Claim 57; SEQ ID NO 263; 408pp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a cell proliferation-related polypeptide. The nucleic acid
CC molecule and the encoded polypeptide, and methods are useful for
CC modulating cell proliferation, senescence, differentiation, development,
CC and stress response in plants, and for producing enhanced food crops. The
CC present sequence represents a cell proliferation-related nucleic acid
CC sequence. The present sequence is published separately from the main body
CC of the specification as EPO data.
XX
SQ Sequence 1929 BP; 575 A; 459 C; 463 G; 432 T; 0 U; 0 Other;

Query Match 10.2%; Score 278.8; DB 12; Length 1929;
Best Local Similarity 59.0%; Pred. No. 1.2e-53;
Matches 519; Conservative 0; Mismatches 352; Indels 9; Gaps 2;

```
Qy        765 CTCTAAGTATCTGAAAGCTGCACAAGAGCTTCTTGATGAAGTTGTTAATATTGTTGGAAA 824
              ||| |||||||||||||| ||||||||  | ||||||||  || || | | ||      |
Db        588 CTCAAAGTATCTGAAAGCAGCACAAGAATTGCTTGATGAGGTCGTCAGTGTTTGGAAGAG 647

Qy        825 AAGCATCAAAGGAGATGATCAAAAGAAGGATAATTCAATGAATAAAGAATCAATGCCTTT 884
              |  |   |    | | |  || |  ||| | |  ||   ||   |||    |
Db        648 TATAAAGCAGAAGGCTCAGAAAGACCAGGCTGAAGCAGGAAAATCAGATAACAAAGAAGC 707

Qy        885 GGCTAGTGATGTCAACACTAATAGTTCTGGTGGTGGTGAAAGTAGCAGCAGGCAGAAAAA 944
              |   | | |       | ||  ||   || |     |   |          |   |
Db        708 CGAGGGGGGT------TCGAAAGGTGAGGGGGTATCTTCCAACCCACAGGAGTCTACTGC 761

Qy        945 TGAAGTTGCTGTTGAGCTTACAACTGCTCAAAGACAAGAACTTCAAATGAAAAAAGCCAA 1004
                | | |||    ||| || |  ||||| | | ||||||| || || |   || |  || ||
Db        762 CAATGCTGCACCAGAGATTTCTGCTGCTGAGAAACAAGAGCTCCAGAATAAGATGGCAAA 821

Qy       1005 GCTTCTTGCCATGCTTGAAGAGGTGGAGCAAAGGTACAGACAGTACCATCACCAAATGCA 1064
               ||| | ||||||  | || ||||||||  | || |||||  ||| ||  |||| ||||||||
Db        822 ACTTATGGCCATGTTGGATGAGGTGGACCGAAAGTACAAACACTATTATCATCAAATGCA 881

Qy       1065 AATAATTGTATTATCATTTGAGCAAGTAGCAGGAATTGGATCAGCCAAATCATACACTCA 1124
               |||  | ||  | |||  |||||     || || ||    ||||  | |||||   | || |||
Db        882 AATTGTAGTCTCATCTTTTGATATGGTTGCTGGGTCTGGAGCTGCCAAGCCTTATACTGC 941

Qy       1125 ATTAGCTTTGCATGCAATTTCGAAGCAATTCAGATGCCTAAAGGATGCAATTGCTGAGCA 1184
              | | ||   | ||  |||| || || || |||||||| || || ||||| ||    || ||
Db        942 AGTGGCCCTTCAGACAATCTCAAAACATTTCAGATGTCTGAAAGATGCTATCAACGATCA 1001

Qy       1185 AGTAAAGGCGACGAGCAAGAGTTTAGGTGAAGAGGAAGGCTTGGGAGGGAAAATCGAAGG 1244
              | || |  |   | ||||   | || || |||||  | |      || |||   | |
Db       1002 GATCAATGTTATCCGGAAGAAACTTGGAGAGGAGGAGAGTTCATCTGGCAAAGAGGGAAA 1061

Qy       1245 CTCAA---GACTCAAATTTGTGGACCATCATCTAAGGCAACAACGCGCGCTGCAACAGAT 1301
               | ||   | |||   | | | ||||| ||  |||| ||||| |||||  | || |||
Db       1062 ATTAACGCGCCTCCGTTATATTGACCAGCAATTAAGACAACAGCGCGCTTTCCAGCAGTA 1121

Qy       1302 AGGAATGATGCAACCAAATGCTTGGAGACCCCAAAGAGGTTTACCTGAAAGAGCTGTCTC 1361
               ||  || | || | ||||||||||||| || || || ||  | || ||||   | || ||
Db       1122 TGGTTTGTTACAGCAAAATGCTTGGAGGCCACAGAGGGGACTGCCCGAAAACTCAGTTTC 1181

Qy       1362 TGTCCTTCGTGCTTGGCTTTTCGAGCATTTTCTTCATCCTTACCCAAAGGATTCAGACAA 1421
               | |||||||||||||||||| || || || ||||| || || ||||| ||||||||| ||
Db       1182 AATTCTTCGTGCTTGGCTTTTTGAACACTTCCTTCACCCGTATCCAAAAGATTCAGAAAA 1241

Qy       1422 AATCATGCTTGCTAAGCAAACGGGGCTAACAAGGAGCCAGGTGTCTAACTGGTTCATAAA 1481
               | ||||| || |  |||| ||  ||||||| || ||| | || || ||||||||||||
Db       1242 GCTGATGCTAGCGAGACAAACTGGCTTAACAAGAAGTCAGATTTCAAATTGGTTCATAAA 1301

Qy       1482 TGCTCGAGTTCGATTATGGAAGCCAATGGTAGAAGAAATGTACTTGGAAGAAGTGAAGAA 1541
              ||| || || ||  | ||||| || ||| | ||||| |||||    |||||  |   | |
Db       1302 TGCCCGTGTCCGCCTGTGGAAACCGATGATCGAAGACATGTATAAAGAAGAGATTGGGGA 1361

Qy       1542 TCAAGAACAAAACAGTACTAATACTTCAGGAGATAACAAAAACAAAGAGACCAATATAAG 1601
                || |  ||  |      | || |   |      ||  |   ||  ||| |||
Db       1362 GGCGGATCTCGACTCGAACTCCTCCTCCGACAACGTACCAAGGAGCAAGGACAAAATAGC 1421

Qy       1602 TGCTCCAAATGAAGAGAAACATCCAATTATTACTAGCAGC 1641
                | |  | ||  || || |||  |  |   |||  |||
Db       1422 AACATCTGAAGATAAGGAAGATCTGAAAAGCTCTATGAGC 1461
```

RESULT 6
AAC36460
ID AAC36460 standard; DNA; 2031 BP.
XX
AC AAC36460;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 13888.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.

PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.

PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 9.5%; Score 261; DB 3; Length 2031;
Best Local Similarity 66.4%; Pred. No. 1.5e-49;
Matches 407; Conservative 0; Mismatches 200; Indels 6; Gaps 2;

```
Qy       951 TGCTGTTGAGCTTACAACTGCTCAAAGACAAGAACTTCAAATGAAAAAAGCCAAGCTTCT 1010
            | ||| |||  |  |    ||  | || ||||||| | || |  ||      | || ||| |
Db      1023 TACTGCTGAAATCCCTCAAGCAGAGAGGCAAGAATTGCAGAGCAAGTTGTCGAAACTTTT 1082

Qy      1011 TGCCATGCTTGAAGAGGTGGAGCAAAGGTACAGACAGTACCATCACCAAATGCAAATAAT 1070
              | ||  | || ||||||||    ||  ||||  |||||  | || || ||||| ||| |
Db      1083 ATCGATATTAGACGAGGTGGATAGAAATTACAAGCAGTATTACCATCAGATGCAGATAGT 1142

Qy      1071 TGTATTATCATTTGAGCAAGTAGCAGGAATTGGATCAGCCAAATCATACACTCAATTAGC 1130
             ||| |  || || ||   | |||| |||  |||| |||||||| |||||||     | ||
Db      1143 TGTGTCGTCTTTCGATGTAATAGCCGGATGTGGAGCAGCCAAACCATACACAGCCCTTGC 1202

Qy      1131 TTTGCATGCAATTTCGAAGCAATTCAGATGCCTAAAGGATGCAATTGCTGAGCAAGTAAA 1190
              | ||  | || |||| ||| ||| | ||| ||| |||||| ||  | |  ||| |
Db      1203 GCTTCAGACCATCTCGAGGCATTTCCGTTGCTTAAGGGATGCGATATCCGGACAAATCTT 1262

Qy      1191 GGCGACGAGCAAGAGTTT---AGGTGAAGAGGAAGGCTTGGGAGGGAAAATCGAAGGCTC 1247
             || ||  || || |||||    ||| ||| |||| || |  |  |||| |    |  ||
Db      1263 GGTGATAAGGAAAAGTTTAGGAGGGGAACAGGATGGATCAGATGGGAGAGGAGTGGGGAT 1322

Qy      1248 A---AGACTCAAATTTGTGGACCATCATCTAAGGCAACAACGCGCGCTGCAACAGATAGG 1304
             |   ||  | |    ||| || || ||  |||||||||| | ||| |||| | | ||||
Db      1323 AAGCAGGTTAAGGAATGTTGATCAACAGGTAAGGCAACAAAGAGCGTTGCAGCGGTTAGG 1382

Qy      1305 AATGATGCAACCAAATGCTTGGAGACCCCAAAGAGGTTTACCTGAAAGAGCTGTCTCTGT 1364
              ||||||||||  |  ||||| | || ||| | |||||||| ||      |||| |  ||
Db      1383 TGTGATGCAACCTCACACTTGGCGGCCTCAACGCGGTTTACCAGATTCTTCTGTTTTGGT 1442

Qy      1365 CCTTCGTGCTTGGCTTTTCGAGCATTTTCTTCATCCTTACCCAAAGGATTCAGACAAAAT 1424
              || |||||||||| || |||||||| || || ||||| ||||||||||||||||| ||
Db      1443 TCTCCGTGCTTGGCTATTTGAGCATTTCCTCCACCCTTATCCAAAGGATTCAGACAAGAT 1502

Qy      1425 CATGCTTGCTAAGCAAACGGGGCTAACAAGGAGCCAGGTGTCTAACTGGTTCATAAATGC 1484
             |||||| ||||  ||||||||| | |   |  ||||||| || ||||||||||||||||
Db      1503 CATGCTCGCTAGACAAACGGGGTTGAGCCGAGGCCAGGTATCGAACTGGTTCATAAATGC 1562

Qy      1485 TCGAGTTCGATTATGGAAGCCAATGGTAGAAGAAATGTACTTGGAAGAAGTGAAGAATCA 1544
              || || ||  | ||||| || ||||| || || ||||||  ||| ||| | |   ||
Db      1563 GCGTGTGCGTCTCTGGAAACCGATGGTGGAGGAGATGTACAAGGAGGAATTCACAGATGC 1622

Qy      1545 AGAACAAAACAGT 1557
             |   ||  | | |
Db      1623 ATTGCAGGAGAAT 1635
```

RESULT 7
AAC43406
ID AAC43406 standard; DNA; 1449 BP.
XX
AC AAC43406;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 39140.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.

PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 9.5%; Score 259.4; DB 3; Length 1449;
Best Local Similarity 66.2%; Pred. No. 3.2e-49;
Matches 406; Conservative 0; Mismatches 201; Indels 6; Gaps 2;

```
Db        717 GGTGATAAGGAAAAGTTTAGGAGGGGAACAGGATGGATCAGATGGGAGAGGAGTGGGGAT 776

Qy       1248 A---AGACTCAAATTTGTGGACCATCATCTAAGGCAACAACGCGCGCTGCAACAGATAGG 1304
              |   ||  | |    ||| || || ||  |||||||||| | ||| |||| | | ||||
Db        777 AAGCAGGTTAAGGAATGTTGATCAACAGGTAAGGCAACAAAGAGCGTTGCAGCGGTTAGG 836

Qy       1305 AATGATGCAACCAAATGCTTGGAGACCCCAAAGAGGTTTACCTGAAAGAGCTGTCTCTGT 1364
               ||||||||||  |  ||||| | || ||| | |||||||| ||     |||| |  ||
Db        837 TGTGATGCAACCTCACACTTGGCGGCCTCAACGCGGTTTACCAGATTCTTCTGTTTTGGT 896

Qy       1365 CCTTCGTGCTTGGCTTTTCGAGCATTTTCTTCATCCTTACCCAAAGGATTCAGACAAAAT 1424
               || |||||||||||  || |||||||| || || ||||| ||||||||||||||||| ||
Db        897 TCTCCGTGCTTGGCTATTTGAGCATTTCCTCCACCCTTATCCAAAGGATTCAGACAAGAT 956

Qy       1425 CATGCTTGCTAAGCAAACGGGGCTAACAAGGAGCCAGGTGTCTAACTGGTTCATAAATGC 1484
              ||||||  ||||  |||||||||| | |    |  ||||||| || |||||||||||||||
Db        957 CATGCTCGCTAGACAAACGGGGTTGAGCCGAGGCCAGGTATCGAACTGGTTCATAAATGC 1016

Qy       1485 TCGAGTTCGATTATGGAAGCCAATGGTAGAAGAAATGTACTTGGAAGAAGTGAAGAATCA 1544
               || || ||  | ||||| || ||||| || || ||||||  ||| ||| | |   ||
Db       1017 GCGTGTGCGTCTCTGGAAACCGATGGTGGAGGAGATGTACAAGGAGGAATTCACAGATGC 1076

Qy       1545 AGAACAAAACAGT 1557
              |   ||  | | |
Db       1077 ATTGCAGGAGAAT 1089
```

RESULT 8
ADG88236
ID ADG88236 standard; cDNA; 1449 BP.
XX
AC ADG88236;
XX
DT 22-APR-2004 (first entry)
XX
DE A. thaliana RPP4-upregulated pathogen infection-related gene #678.
XX
KW Pathogen infection-related gene; plant; Peronospora parasitica;
KW defence mechanism; RPP4; pathogen resistance; transgenic plant; oomycete;
KW fungus; bacterium; virus; nematode; insect; aphid; gene; ss.
XX
OS Arabidopsis thaliana.
XX
PN WO200222675-A2.
XX
PD 21-MAR-2002.
XX
PF 14-SEP-2001; 2001WO-US028506.
XX
PR 15-SEP-2000; 2000US-0232778P.
PR 22-JUN-2001; 2001US-0300183P.
XX
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
PA (UYNC-) UNIV NORTH CAROLINA.
PA (GLAZ/) GLAZEBROOK J.
PA (WANG/) WANG X.
PA (DANG/) DANGL J L.
PA (EULG/) EULGEM T.
PA (ZHUT/) ZHU T.
XX
PI Glazebrook J, Wang X, Dangl JL, Eulgem T, Zhu T;
XX
DR WPI; 2002-292409/33.
XX
PT Novel isolated polynucleotide, useful for conveying pathogen resistance
PT to plants, and for identifying plants infected with a pathogen.
XX
PS Claim 3; SEQ ID NO 678; 605pp; English.
XX
CC The invention relates to 691 Arabidopsis thaliana genes (ADG87559--
CC ADG87557}) whose expression is altered in response to pathogen infection,
CC and to homologues of these genes from other plants or fungi, especially
CC from maize, soybean, barley, alfalfa, sunflower, canola (oilseed rape),
CC cotton, peanut, sorghum, tobacco, sugarbeet, rice or wheat. The
CC expression of genes of the invention was upregulated or downregulated in
CC Arabidopsis plants infected with the oomycete Peronospora parasitica,
CC indicating that they play a role in defence mechanisms. The genes of the
CC invention are regulated by RPP7 or RRP8 which act via unconventional
CC signalling cascades, or by the RPP4-dependent pathway. The invention also
CC relates to polypeptides encoded by the pathogen infection-related genes;
CC promoter motifs from pathogen infection-related genes (ADG88243-ADG88327)
CC ; expression cassettes, host cells and pathogen-resistant transgenic
CC plants and their progeny comprising a polynucleotide of the invention;
CC and a method of identifying a plant cell infected with a pathogen. The
CC polynucleotide sequences and methods of the invention are useful for
CC identifying plants infected with a pathogen, and for conferring
CC resistance to pathogens such as oomycetes, fungi, bacteria, viruses,
CC neamotodes and insects (e.g., aphids). The present sequence represents an
CC Arabidopsis thaliana gene whose expression is altered in response to
CC Peronospora parasitica infection. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1449 BP; 437 A; 289 C; 365 G; 358 T; 0 U; 0 Other;

Query Match 9.5%; Score 259.4; DB 6; Length 1449;
Best Local Similarity 66.2%; Pred. No. 3.2e-49;
Matches 406; Conservative 0; Mismatches 201; Indels 6; Gaps 2;

```
Qy        951 TGCTGTTGAGCTTACAACTGCTCAAAGACAAGAACTTCAAATGAAAAAAGCCAAGCTTCT 1010
              | ||| |||  |  |    ||  | || |||||| | || |  ||     | || ||| |
Db        477 TACTGCTGAAATCCCTCAAGCAGAGAGGCAAGAATTGCAGAGCAAGTTGTCGAAACTTTT 536

Qy       1011 TGCCATGCTTGAAGAGGTGGAGCAAAGGTACAGACAGTACCATCACCAAATGCAAATAAT 1070
               | ||  | || ||||||||   ||  ||||  |||||  | || || ||||| ||| |
Db        537 ATCGATATTAGACGAGGTGGATAGAAATTACAAGCAGTATTACCATCAGATGCAGATAGT 596

Qy       1071 TGTATTATCATTTGAGCAAGTAGCAGGAATTGGATCAGCCAAATCATACACTCAATTAGC 1130
               || |  || || ||   | |||| |||  |||| ||||||||  |||||||     | ||
Db        597 CGTGTCGTCTTTCGATGTAATAGCCGGATGTGGAGCAGCCAAACCATACACAGCCCTTGC 656

Qy       1131 TTTGCATGCAATTTCGAAGCAATTCAGATGCCTAAAGGATGCAATTGCTGAGCAAGTAAA 1190
               | ||  | || |||| ||| ||| | ||| ||| |||||| ||  | |  ||| |
Db        657 GCTTCAGACCATCTCGAGGCATTTCCGTTGCTTAAGGGATGCGATATCCGGACAAATCTT 716

Qy       1191 GGCGACGAGCAAGAGTTT---AGGTGAAGAGGAAGGCTTGGGAGGGAAAATCGAAGGCTC 1247
              || ||  || || |||||   ||| ||| |||| || |  |  |||| |   |  ||
Db        717 GGTGATAAGGAAAAGTTTAGGAGGGGAACAGGATGGATCAGATGGGAGAGGAGTGGGGAT 776

Qy       1248 A---AGACTCAAATTTGTGGACCATCATCTAAGGCAACAACGCGCGCTGCAACAGATAGG 1304
              |   ||  | |    ||| || || ||  |||||||||| | ||| |||| | | ||||
Db        777 AAGCAGGTTAAGGAATGTTGATCAACAGGTAAGGCAACAAAGAGCGTTGCAGCGGTTAGG 836

Qy       1305 AATGATGCAACCAAATGCTTGGAGACCCCAAAGAGGTTTACCTGAAAGAGCTGTCTCTGT 1364
               ||||||||||  |  ||||| | || ||| | |||||||| ||     |||| |  ||
Db        837 TGTGATGCAACCTCACACTTGGCGGCCTCAACGCGGTTTACCAGATTCTTCTGTTTTGGT 896

Qy       1365 CCTTCGTGCTTGGCTTTTCGAGCATTTTCTTCATCCTTACCCAAAGGATTCAGACAAAAT 1424
               || |||||||||||  || |||||||| || || ||||| ||||||||||||||||| ||
Db        897 TCTCCGTGCTTGGCTATTTGAGCATTTCCTCCACCCTTATCCAAAGGATTCAGACAAGAT 956

Qy       1425 CATGCTTGCTAAGCAAACGGGGCTAACAAGGAGCCAGGTGTCTAACTGGTTCATAAATGC 1484
              ||||||  ||||  |||||||||| | |    |  ||||||| || |||||||||||||||
Db        957 CATGCTCGCTAGACAAACGGGGTTGAGCCGAGGCCAGGTATCGAACTGGTTCATAAATGC 1016

Qy       1485 TCGAGTTCGATTATGGAAGCCAATGGTAGAAGAAATGTACTTGGAAGAAGTGAAGAATCA 1544
               || || ||  | ||||| || ||||| || || ||||||  ||| ||| | |   ||
Db       1017 GCGTGTGCGTCTCTGGAAACCGATGGTGGAGGAGATGTACAAGGAGGAATTCACAGATGC 1076

Qy       1545 AGAACAAAACAGT 1557
              |   ||  | | |
Db       1077 ATTGCAGGAGAAT 1089
```

RESULT 9

ADO62706
ID ADO62706 standard; DNA; 1971 BP.
XX
AC ADO62706;
XX
DT 15-JUL-2004 (first entry)
XX
DE Transcription factor G2550 orthologous sequence, SEQ ID 1173.
XX
KW Plant; transcription factor; transgenic plant; abiotic stress tolerance;
KW osmotic stress tolerance; cold tolerance; heat tolerance;
KW low nitrogen tolerance; low phosphate tolerance; fungal disease;
KW glyphosate resistance; flowering; fertility; seed development; ds.
XX
OS Glycine max.
XX
PN WO2004031349-A2.
XX
PD 15-APR-2004.
XX
PF 18-SEP-2003; 2003WO-US030292.
XX
PR 18-SEP-2002; 2002US-0411837P.
PR 17-DEC-2002; 2002US-0434166P.
PR 24-APR-2003; 2003US-0465809P.
XX
PA (MEND-) MENDEL BIOTECHNOLOGY INC.
XX
PI Jiang C, Heard JE, Ratcliffe O, Creelman RA, Adam LJ, Reuber TL;
PI Riechmann JL, Haake V, Dubell AN, Keddie JS, Sherman BK;
XX
DR WPI; 2004-330163/30.
XX
PT New recombinant polynucleotide encoding transcription factor
PT polypeptides, useful for producing transgenic plants with advantageous
PT properties compared to a reference plant.
XX
PS Claim 1; SEQ ID NO 1173; 510pp; English.
XX
CC The present invention relates to novel plant transcription factor
CC proteins (I) and nucleotide sequences (II) (ADO61534-ADO63778). The
CC sequences can be used to produce transgenic plants, which overexpress
CC (II), where the transgenic plant has an altered trait as compared to a
CC non-transgenic plant or wild-type plant. The transgenic plant comprises
CC an altered trait selected from increased tolerance to abiotic stress,
CC increased tolerance to osmotic stress, increased tolerance to cold,
CC increased germination in cold, increased tolerance to heat, increased
CC germination in heat, increased tolerance to freezing conditions,
CC increased tolerance to low nitrogen conditions, increased tolerance to
CC low phosphate conditions, increased tolerance to disease, including
CC fungal disease and particularly Erysiphe, Fusarium and Botrytis,
CC increased tolerance to multiple fungal pathogens, increased resistance to
CC glyphosate, increased sensitivity to ABA, reduced sensitivity to ABA,
CC increased sensitivity to ACC, altered sugar sensing, increased tolerance
CC to sugars, altered carbon/nitrogen sensing, early flowering, late
CC flowering, altered flower structure, loss of flower determinacy, reduced
CC fertility, altered shoot meristem development, altered branching pattern,
CC altered stem morphology, altered vascular tissue structure, reduced
CC apical dominance, altered trichome density, altered trichome development,
CC altered trichome structure, altered root development, altered shade
CC avoidance, altered seed development, altered seed ripening, altered seed
CC germination, slow growth, fast growth, altered cell differentiation,
CC altered cell proliferation, altered cell expansion, altered phase change,
CC altered senescence, abnormal embryo development, altered programmed cell
CC death, lethality when overexpressed, altered necrosis patterns, increased
CC plant size, increased biomass, large seedlings, dwarfed plants, dark
CC green leaves, change in leaf shape, increased leaf size and mass, light
CC green or gray leaves, glossy leaves, altered abaxial/adaxial polarity,
CC altered seed coloration, altered seed size, altered seed shape, large
CC seed, increased leaf wax, increased leaf fatty acids, altered seed oil
CC content, altered seed protein content, altered seedprenyl content,
CC altered leaf prenyl lipid content, increased anthocyanin levels, and
CC decreased anthocyanin levels. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1971 BP; 667 A; 360 C; 451 G; 493 T; 0 U; 0 Other;

Query Match 8.9%; Score 244.4; DB 12; Length 1971;
Best Local Similarity 57.7%; Pred. No. 9.9e-46;
Matches 512; Conservative 0; Mismatches 346; Indels 30; Gaps 3;

```
Qy       741 AAGTTCTAGCAACATGGTTTTAGGCTCTAAGTATCTGAAAGCTGCACAAGAGCTTCTTGA 800
Db       201 AGGCTCGAGCAACAATGTCTTGAACTCACAATACCTCAAGGCAGCACAGGAGTTGCTTGA 260

Qy       801 TGAAGTTGTTAATATTGTTGGAAAAAGCATCAAAGGAGATGATCAAAAGAAGGATAATTC 860
Db       261 TGAAATAGTAAATGTCCGAAAGGCTTTGAAGCAAACTGGTTTGGAAAAGCAACAGAGTTT 320

Qy       861 AATGAATA---------AAGAATCAATGCCTTTGGCTAGTGATGTCAACACTAATAGTTC 911
Db       321 CCGTGACACTGGTTTAGATGGCTCCAAAGATTCTGATGGAAAATCTACCAGCCAATCTGT 380

Qy       912 TGGTGGTGGTGAAAGTAGCAGCAGGCAGAAAAATGAAGTTGCTGTTGAGCTTACAACTGC 971
Db       381 GCAGATGTCTTCAGGCCCCAATGGTTCCGCTGCTAACGCTTCTTGTGAGCTATCACCTGC 440

Qy       972 TCAAAGACAAGAACTTCAAATGAAAAAAGCCAAGCTTCTTGCCATGCTTGAAGAGGTGGA 1031
Db       441 AGAACGGCAGAACTTGTTGGACAAGAAGACAAAGCTTTTGTCCATGCTGGATGAGGTGGA 500

Qy       1032 GCAAAGGTACAGACAGTACCATCACCAAATGCAAATAATTGTATTATCATTTGAGCAAGT 1091
Db       501 TAAAAGATACAGACAGTACTGCCATCAGATGCAGATTGTGGTGTCATCTTTTGACATGGT 560

Qy       1092 AGCAGGAATTGGATCAGCCAAATCATACACTCAATTAGCTTTGCATGCAATTTCGAAGCA 1151
Db       561 TGCTGGCTGTGGAGCAGCAGAACCATATACCACACTTGCCTTAAGAACAATTTCTCGCCA 620

Qy       1152 ATTCAGATGCCTAAAGGATGCAATTGCTGAGCAAGTAAAGGCGACGAGCAAGAGTTTAGG 1211
Db       621 CTTTCGGTGTTTGCGTGATGCCATCAGTGGCCAAATTCAGGTGACCCAAAGAAGCCTTGG 680

Qy       1212 TGAAGAGGAAGGCTTGGGAGGGAAAATCGAAGGCTCAAGACTCAAATTTGTGGACCATCA 1271
Db       681 GGAGCAAGAGGG----------AATACCCCGTCTCCGCTATGTGGATCAGCA 722

Qy       1272 TCTAAGGCAACAACGCGCGCTGCAACAGATAGGAATGATGCAACCAAATGCTTGGAGACC 1331
Db       723 ACTTAGACAACAAAAGGCCCTTCAGCAACTTGGTGTAATGAGAC---AAGCTTGGAGGCC 779

Qy       1332 CCAAAGAGGTTTACCTGAAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTTTTCGAGCATTT 1391
Db       780 TCAGAGGGGACTTCCTGAAACCTCTGTTTCAGTACTCCGTGCTTGGCTCTTTGAGCATTT 839

Qy       1392 TCTTCATCCTTACCCAAAGGATTCAGACAAAATCATGCTTGCTAAGCAAACGGGGCTAAC 1451
Db       840 CCTTCATCCTTATCCTAAGGATTCAGAGAAAATTATGCTAGCAAGGCAAACTGGCTTAAC 899

Qy       1452 AAGGAGCCAGGTGTCTAACTGGTTCATAAATGCTCGAGTTCGATTATGGAAGCCAATGGT 1511
Db       900 AAGAAACCAGGTGGCAAACTGGTTCATTAATGCAAGGGTGCGTCTATGGAAGCCAATGGT 959

Qy       1512 AGAAGAAATGTACTTGGAAGAAGTGAAGAATCAAGAACAAAACAGTACTAATACTTCAGG 1571
Db       960 TGAGGAAATGTACAAAGAAGAATTTGGTGATTCTGAGATGAGCAGCAATCTATTATCATC 1019

Qy       1572 AGATAACAAAAACAAAGAGACCAATATAAGTGCTCCAAATGAAGAGAA 1619
Db       1020 AGAGAAAACACTCAAAGCTCCAAGAGATGATGTTCAAGCTTCTGACAA 1067
```

RESULT 10
ABK82114

ID ABK82114 standard; cDNA; 706 BP.
XX
AC ABK82114;
XX
DT 27-AUG-2002 (first entry)
XX
DE DNA encoding novel floral meristem identity protein LpHBb.
XX
KW Ryegrass; fescue; MADS-box; MADS; MADS-like protein; CENTRORADIALIS; CEN;
KW CEN-like protein; APETALA2; AP2; AP2-like protein; HB; Homeo-box protein;
KW HB-like protein; plant growth; plant architecture;
KW inflorescence development; flower development; embryo development;
KW seed development; flower organ identity; phase change; male sterility;
KW hybrid seed production; herbage quality; early maturing crop;
KW biomass increase; branching increase; blocking flowering;
KW allergenic pollen; floral meristem identity protein; gene; ss.
XX
OS Lolium perenne.
XX
PN WO200233091-A1.
XX
PD 25-APR-2002.
XX
PF 17-OCT-2001; 2001WO-AU001311.
XX
PR 19-OCT-2000; 2000AU-00000873.
XX
PA (AGRI-) AGRIC VICTORIA SERVICES PTY LTD.
PA (AGRE-) AGRESEARCH LTD.
XX
PI Spangenberg G, Sawbridge TI, Ong EK, Emmerling M;
XX
DR WPI; 2002-454601/48.
DR P-PSDB; ABG60937.
XX
PT New substantially purified or isolated polypeptide e.g., MADS-box,
PT CENTRORADIALIS, APETALA2, Homeo-box proteins, isolated from ryegrass or
PT fescue species, useful for controlling plant life cycles and/or growth
PT phases.
XX
PS Claim 6; Fig 41; 290pp; English.
XX
CC The invention describes a substantially purified or isolated polypeptide
CC (I) from a ryegrass (Lolium sp.) or fescue (Festuca sp.) species, such as
CC MADS-box (MADS) and MADS-like proteins, CENTRORADIALIS (CEN) and CEN-like
CC proteins, APETALA2 (AP2) and AP2-like proteins, Homeo-box proteins (HB)
CC and HB-like proteins, or their functionally active fragments or variants.
CC Nucleic acid (II) encoding (I), a construct (III) comprising (II), or a
CC vector (IV) comprising (II), is useful for modifying plant life cycles
CC and/or growth phases, flowering processes, flowering and/or plant
CC architecture and/or flower and/or inflorescence development in a plant,
CC which involves introducing (II), (III) or (IV) into the plant. The
CC individual or simultaneous enhancement or downregulation of MADS-box gene
CC activities may alter flower, embryo and seed development, e.g., enhance
CC or inhibit embryo differentiation and growth or alter flower organ
CC identity through conversion of one floral organ in another. Manipulation
CC of CEN or HB activity in plants alters the control of phase change,
CC flowering time and the number of leaves made before flowering.
CC Manipulation of AP2 activity in plants alters flower organ identity
CC through conversion of one floral organ in another, leads to a change in
CC the number of floral organs and flowering architecture. Manipulation of
CC flowering plant architecture has a wide range of applications such as:
CC inducing male sterility for hybrid seed production; changing flower
CC architecture for enhancing value of ornamentals; delaying flowering in
CC forage grasses thus stopping the formation of less digestible stems and
CC increasing herbage quality; altering flowering time allowing early
CC maturing crops; delaying vegetative phase to increase biomass production;
CC increasing branching to enhanced business in fruit trees; altering plant
CC size e.g. shorter plant stature; and in blocking flowering and reducing
CC release for allergenic pollen. This sequence encodes a novel floral
CC meristem identity protein involved in floral development and a potential
CC target for manipulating plant life cycles
XX

SQ Sequence 706 BP; 210 A; 148 C; 188 G; 147 T; 0 U; 13 Other;

Query Match 8.7%; Score 237.6; DB 6; Length 706;
Best Local Similarity 66.7%; Pred. No. 2.5e-44;
Matches 355; Conservative 0; Mismatches 174; Indels 3; Gaps 1;

```
Qy       947 AAGTTGCTGTTGAGCTTACAACTGCTCAAAGACAAGAACTTCAAATGAAAAAAGCCAAGC 1006
                | | ||    ||||||| | |||||| | |  ||||||||||| |  || |  ||||| |
Db       173 ATGCTGTACCTGAGCTTTCCACTGCTGAGAAGCAAGAACTTCAGAACAAGATGGCCAAAC 232

Qy      1007 TTCTTGCCATGCTTGAAGAGGTGGAGCAAAGGTACAGACAGTACCATCACCAAATGCAAA 1066
               |  | || ||| | || |||||||| |  |  ||||  || ||  | ||||||||||||
Db       233 TGATGGCAATGTTGGATGAGGTGGACCGGAAATACAAGCACTATTACCACCAAATGCAAA 292

Qy      1067 TAATTGTATTATCATTTGAGCAAGTAGCAGGAATTGGATCAGCCAAATCATACACTCAAT 1126
               | || | ||| |||||     || || ||   |||||| || ||  | ||||||  |
Db       293 ATGTGGTTTCATCTTTTGATGTGGTGGCTGGGCCTGGATCTGCAAAGCCGTACACTGCAG 352

Qy      1127 TAGCTTTGCATGCAATTTCGAAGCAATTCAGATGCCTAAAGGATGCAATTGCTGAGCAAG 1186
               | ||| | ||  |||| |||  ||| ||| | ||| | |||||||| ||   ||| ||
Db       353 TTGCTCTTCAGACAATCTCGCGGCACTTCCGGTGCTTGAAGGATGCCATCAATGATCAGA 412

Qy      1187 TAAAGGCGACGAGCAAGAGTTTAGGT---GAAGAGGAAGGCTTGGGAGGGAAAATCGAAG 1243
               | ||  |  | || ||||   | |||   |||||| |     |  ||   | ||  | ||
Db       413 TTAATGTTATCAGGAAGAAGCTTGGTGAGGAAGAGAATTCATCTGGCAAGGAAGGCAAAT 472

Qy      1244 GCTCAAGACTCAAATTTGTGGACCATCATCTAAGGCAACAACGCGCGCTGCAACAGATAG 1303
                 |  |||| |   | || || || || ||||| || || ||  | ||||||  |
Db       473 TAACCCGTCTCAGGTACATCGATCAGCAGCTGAGGCAGCAGCGAGCTTTCCAACAGTATG 532

Qy      1304 GAATGATGCAACCAAATGCTTGGAGACCCCAAAGAGGTTTACCTGAAAGAGCTGTCTCTG 1363
               | |||||  |  | ||| || |||||||| |||||||  | |||||||   | ||  | |
Db       533 GCATGATTCCGCAAAACGCATGGAGACCGCAAAGAGGACTGCCTGAAAACTCGGTTACGG 592

Qy      1364 TCCTTCGTGCTTGGCTTTTCGAGCATTTTCTTCATCCTTACCCAAAGGATTCAGACAAAA 1423
               | || |||||||||||||||||| || || ||||| || |||||||| ||||||||  ||
Db       593 TTCTCCGTGCTTGGCTTTTCGAACACTTCCTTCACCCGTACCCAAAAGATTCAGAAAAGT 652

Qy      1424 TCATGCTTGCTAAGCAAACGGGGCTAACAAGGAGCCAGGTGTCTAACTGGTT 1475
               | ||||| || | ||| || ||  | ||||||||||| | || || |||||
Db       653 TGATGCTAGCGAGGCAGACAGGCTTGACAAGGAGCCAGATTTCGAATTGGTT 704
```

RESULT 11
AAC39138
ID AAC39138 standard; DNA; 2025 BP.
XX
AC AAC39138;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 23512.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.

```
PR   13-OCT-1999;   99US-0159295P.
PR   14-OCT-1999;   99US-0159329P.
PR   14-OCT-1999;   99US-0159330P.
PR   14-OCT-1999;   99US-0159331P.
PR   14-OCT-1999;   99US-0159637P.
PR   14-OCT-1999;   99US-0159638P.
PR   18-OCT-1999;   99US-0159584P.
PR   21-OCT-1999;   99US-0160741P.
PR   21-OCT-1999;   99US-0160767P.
PR   21-OCT-1999;   99US-0160768P.
PR   21-OCT-1999;   99US-0160770P.
PR   21-OCT-1999;   99US-0160814P.
PR   21-OCT-1999;   99US-0160815P.
PR   22-OCT-1999;   99US-0160980P.
PR   22-OCT-1999;   99US-0160981P.
PR   22-OCT-1999;   99US-0160989P.
PR   25-OCT-1999;   99US-0161404P.
PR   25-OCT-1999;   99US-0161405P.
PR   25-OCT-1999;   99US-0161406P.
PR   26-OCT-1999;   99US-0161359P.
PR   26-OCT-1999;   99US-0161360P.
PR   26-OCT-1999;   99US-0161361P.
PR   28-OCT-1999;   99US-0161920P.
PR   28-OCT-1999;   99US-0161992P.
PR   28-OCT-1999;   99US-0161993P.
PR   29-OCT-1999;   99US-0162142P.

  Query Match             8.6%;  Score 234.6;  DB 3;  Length 2025;
  Best Local Similarity   58.6%;  Pred. No. 1.8e-43;
  Matches  480;  Conservative    0;  Mismatches  309;  Indels   30;  Gaps    3;

Qy        739 GGAAGTTCTAGCAACATGGTTTTAGGCTCTAAGTATCTGAAAGCTGCACAAGAGCTTCTT 798
              ||| | | ||  | ||  ||| |  | |||  |||||| ||| |||| ||  |  | |||
Db        669 GGAGGGTTTATGAGCAGCGTTCTGCGTTCTCGGTATCTTAAACCTGCTCAGAATTTGCTT 728

Qy        799 GATGAAGTTGTTAATATTGTTGGAAAAAGCATCAAAGGAGATGATCAAAAGAAGGATAAT 858
              |||||||| |||| | |      | ||   | | ||      | |  | |||| | |
Db        729 GATGAAGTGGTTAGTGTCAAGAAAGAACTAAACCAAATGAGTAAGAAGAAGATGAA---- 784

Qy        859 TCAATGAATAAAGAATCAATGCCTTTGGCTAGTGATGTCAACACTAATAGTTCTGGTGGT 918
                    | | |||| ||||  | |  | | | | |   |  | |     |  |
Db        785 ------------AGTTAATGACTTTAACAATGGTTCTAAGGAGATAGAAGGAGGAGGCA 831

Qy        919 GGTGAAAGTAGCAGCAGGCAGAAAAATGAAGTTGCTGTTGAGCTTACAACTGCTCAAAGA 978
              |  |  ||     | | | |    |||       |  ||||| |  | |||  | || |
Db        832 GTGGCGAGCTATCGAATG-ATTCGAATGGGAAATCGATTGAGTTATCTACTATTGAACGT 890

Qy        979 CAAGAACTTCAAATGAAAAAAGCCAAGCTTCTTGCCATGCTTGAAGAGGTGGAGCAAAGG 1038
              | || ||||| |  || ||   |||||| |  | ||| |||| ||||| ||  ||||
Db        891 GATGAGCTTCATAACAAGAAGAACAAGCTTTTAACAATGGTTGATGAGGTAGATAAAAGA 950

Qy       1039 TACAGACAGTACCATCACCAAATGCAAATAATTGTATTATCATTTGAGCAAGTAGCAGGA 1098
              || |  || ||  | || |||||| ||  | | |  | ||||||||||  ||||||||||
Db        951 TATAACCAATATTACCATCAAATGGAAGCATTAGCTTCATCATTTGAGATAGTAGCAGGA 1010

Qy       1099 ATTGGATCAGCCAAATCATACACTCAATTAGCTTTGCATGCAATTTCGAAGCAATTCAGA 1158
              |||||||||| ||  | |||||   | | ||| |  |   ||| ||    || ||  |
Db       1011 CTTGGATCAGCTAAGCCTTACACATCAGTTGCTCTCAACAGAATCTCTCGCCATTTTCGT 1070

Qy       1159 TGCCTAAAGGATGCAATTGCTGAGCAAGTAAAGGCGACGAGCAAGAGTTTAGGTGAAGAG 1218
                ||    || |||||   || ||| |  ||      ||  |||   | || ||  |
Db       1071 GCTCTTCGCGACGCAATAAAGGAACAAATTCAGATTGTTAGAGAGAAACTTGGGGAGAAA 1130

Qy       1219 GAAGGCTTGGGAGGGAAAATCGAAGGC----------TCAAGACTCAAATTTGTGGAC 1266
              | |||   |     | |    ||                |||||||||  | | | ||
Db       1131 GGAGGAGAGTCGTTGGATGAGCAACAAGGAGAGAGGATACCAAGACTCAGGTATTTAGAT 1190

Qy       1267 CATCATCTAAGGCAACAACGCGCGCTGCAACAGATAGGAATGATGCAACCAAATGCTTGG 1326
              || |   | || || ||| | ||  |||| ||          |||   | |   ||||||
Db       1191 CAACGGTTGAGACAGCAAAGAGCTTTGCATCAACAGCTTGGAATGGTTCGACCCGCTTGG 1250

Qy       1327 AGACCCCAAAGAGGTTTACCTGAAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTTTTCGAG 1386
              ||||| ||||||||| | |||||||   |||| ||||| || ||||||||||| |||||
```

```
Db       1251 AGACCTCAAAGAGGTCTTCCTGAAAACTCTGTTTCTGTTCTCCGTGCTTGGCTCTTCGAA 1310

Qy       1387 CATTTTCTTCATCCTTACCCAAAGGATTCAGACAAAATCATGCTTGCTAAGCAAACGGGG 1446
              ||||| |||||||| || || || || || || ||||||||||||| || || || ||
Db       1311 CATTTCCTTCATCCATATCCGAAAGAATCTGAGAAAATCATGCTTGCGAAACAGACAGGA 1370

Qy       1447 CTAACAAGGAGCCAGGTGTCTAACTGGTTCATAAATGCTCGAGTTCGATTATGGAAGCCA 1506
              ||  | | || ||||||  |||| |||||||||| ||  |||||||  ||||||| ||
Db       1371 CTGTCGAAGAACCAGGTTGCTAATTGGTTCATAAACGCGAGAGTTCGTCTATGGAAACCG 1430

Qy       1507 ATGGTAGAAGAAATGTACTTGGAAGAAGTGAAGAATCAA 1545
              ||| | ||||| |||||    |||||  |      || ||
Db       1431 ATGATCGAAGAGATGTATAAAGAAGAGTTTGGTGATGAA 1469


RESULT 12
AAD06493
ID   AAD06493 standard; cDNA; 1575 BP.
XX
AC   AAD06493;
XX
DT   10-AUG-2001  (first entry)
XX
DE   Arabidopsis thaliana G418 transcription factor homologue, G2550 cDNA.
XX
KW   Transcription factor; pesticidal; antimicrobial; gene therapy;
KW   pathogen tolerance; trichome structure; callose induction;
KW   phytoalexin induction; plant structure; plant development; ss.
XX
OS   Arabidopsis thaliana.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .1575
FT                   /*tag= a
FT                   /product= "A. thaliana transcription factor homologue"
XX
PN   WO200135726-A1.
XX
PD   25-MAY-2001.
XX
PF   14-NOV-2000; 2000WO-US031418.
XX
PR   17-NOV-1999;   99US-0166228P.
PR   17-APR-2000; 2000US-0197899P.
PR   22-AUG-2000; 2000US-0227439P.
XX
PA   (MEND-) MENDEL BIOTECHNOLOGY INC.
PA   (HEAR/) HEARD J.
PA   (RATC/) RATCLIFFE O.
PA   (CREE/) CREELMAN R.
PA   (JIAN/) JIANG C.
PA   (PINE/) PINEDA O.
PA   (REUB/) REUBER L.
PA   (ADAM/) ADAM L.
XX
PI   Heard J,  Ratcliffe O,  Creelman R,  Jiang C,  Pineda O,  Reuber L;
PI   Adam L;
XX
DR   WPI; 2001-335978/35.
DR   P-PSDB; AAE02524.
XX
PT   Nucleic acids encoding plant transcription factor polypeptides, useful
PT   for altering the pathogen resistance characteristics of plants, e.g.
PT   corn, potato and cotton plants.
XX
PS   Claim 4; Page 110-112; 134pp; English.
XX
CC   The present sequence is a cDNA encoding Arabidopsis thaliana
CC   transcription factor homologue. The transcription factors are used to
CC   alter the structure and developmental characteristics of plants such as
CC   soybean, wheat, corn, potato, cotton, rice, oilseed rape, sunflower,
CC   alfalfa, sugarcane, turf, banana, blackberry, blueberry, strawberry,
CC   raspberry, carrot, cantaloupe, cauliflower, cucumber, coffee, eggplant,
```

CC grapes, mango, lettuce, honeydew, melon, onion, papaya, peppers,
CC pineapple, spinach, squash, sweet corn, tobacco, tomato, peas,
CC watermelon, rosaceous fruits and vegetable brassicas. The transcription
CC factors are specifically useful for modifying traits associated with
CC plant's pathogen tolerance such as alterations in cell wall composition,
CC trichome number or structure, callose induction, phytoalexin induction,
CC and alterations in the cell death response. Transgenic plants expressing
CC these transcription factors are more tolerant to biotrophic or
CC necrotrophic pathogens such as fungi, bacteria, mollicutes, viruses,
CC nematodes and parasitic higher plants. The transcription factors are also
CC used in gene therapy
XX
SQ Sequence 1575 BP; 512 A; 323 C; 323 G; 417 T; 0 U; 0 Other;

Query Match 8.6%; Score 234.4; DB 4; Length 1575;
Best Local Similarity 61.6%; Pred. No. 1.8e-43;
Matches 392; Conservative 0; Mismatches 241; Indels 3; Gaps 1;

```
Qy       958 GAGCTTACAACTGCTCAAAGACAAGAACTTCAAATGAAAAAAGCCAAGCTTCTTGCCATG 1017
Db       670 GAGTTGTCTCCTTCAGAACGTCAGGAGCTACAGAGCAAGAAGAGCAAGCTTTTAACAATG 729

Qy      1018 CTTGAAGAGGTGGAGCAAAGGTACAGACAGTACCATCACCAAATGCAAATAATTGTATTA 1077
Db       730 GTGGATGAGGTAGATAAAAGGTATAACCAATACCATCATCAAATGGAAGCTTTAGCATCG 789

Qy      1078 TCATTTGAGCAAGTAGCAGGAATTGGATCAGCCAAATCATACACTCAATTAGCTTTGCAT 1137
Db       790 TCTTTCGAGATGGTAACAGGTCTTGGAGCAGCTAAGCCTTACACATCCGTAGCTCTGAAT 849

Qy      1138 GCAATTTCGAAGCAATTCAGATGCCTAAAGGATGCAATTGCTGAGCAAGTAAAGGCGACG 1197
Db       850 AGAATCTCTCGCCATTTCCGCTGTTTACGCGACGCGATAAAAGAACAGATTCAGGTGATC 909

Qy      1198 AGCAAGAGTTTAGGTGAAGAGGAAGGCTTGGGAGGGAAAATCGAAGG---CTCAAGACTC 1254
Db       910 AGAGGGAAGCTTGGGGAGAGAGAGACTTCTGATGAACAAGGAGAGAGGATACCGCGTCTT 969

Qy      1255 AAATTTGTGGACCATCATCTAAGGCAACAACGCGCGCTGCAACAGATAGGAATGATGCAA 1314
Db       970 AGGTACTTAGATCAACGGTTGAGACAACAGAGAGCTTTGCATCAACAACTTGGAATGGTT 1029

Qy      1315 CCAAATGCTTGGAGACCCCAAAGAGGTTTACCTGAAAGAGCTGTCTCTGTCCTTCGTGCT 1374
Db      1030 AGACCAGCTTGGAGACCACAAAGAGGCTTACCTGAAAACTCTGTCTCTATACTTCGAGCT 1089

Qy      1375 TGGCTTTTCGAGCATTTTCTTCATCCTTACCCAAAGGATTCAGACAAAATCATGCTTGCT 1434
Db      1090 TGGCTCTTTGAGCATTTCCTTCATCCATATCCTAAAGAATCAGAGAAAATCATGCTTTCA 1149

Qy      1435 AAGCAAACGGGGCTAACAAGGAGCCAGGTGTCTAACTGGTTCATAAATGCTCGAGTTCGA 1494
Db      1150 AAGCAGACAGGACTATCGAAAAACCAGGTTGCAAATTGGTTTATTAACGCGAGAGTTCGA 1209

Qy      1495 TTATGGAAGCCAATGGTAGAAGAAATGTACTTGGAAGAAGTGAAGAATCAAGAACAAAAC 1554
Db      1210 CTATGGAAACCAATGATTGAAGAGATGTATAAAGAAGAGTTTGGAGAATCAGCAGAGTTA 1269

Qy      1555 AGTACTAATACTTCAGGAGATAACAAAAACAAAGAG 1590
Db      1270 CTCTCTAACTCTAATCAAGACACCAAAAAAATGCAG 1305
```

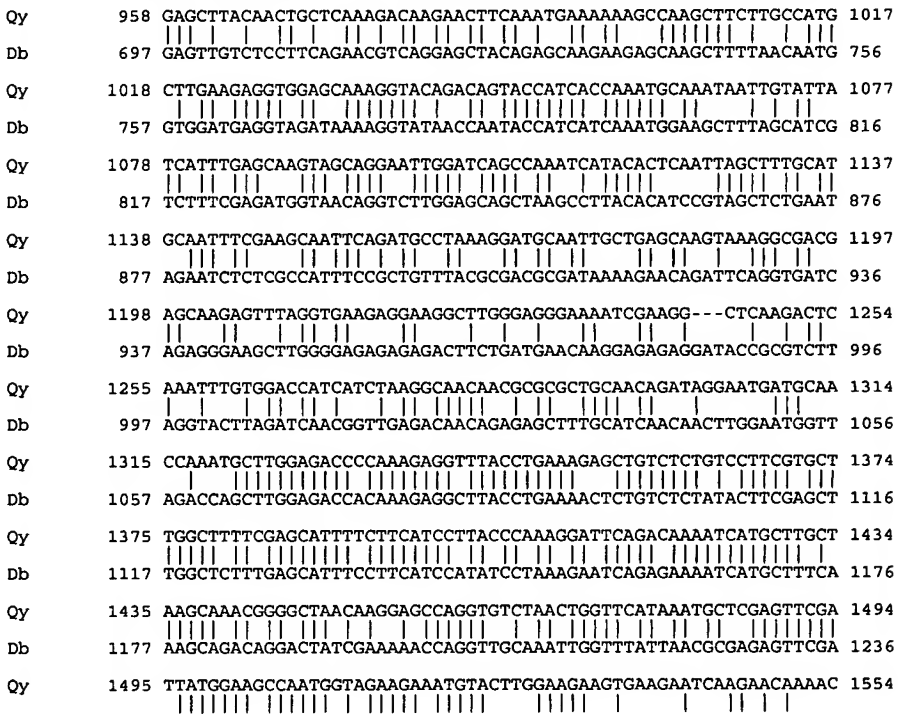
RESULT 13
ADO61968
ID ADO61968 standard; DNA; 1575 BP.
XX
AC ADO61968;
XX
DT 15-JUL-2004 (first entry)
XX
DE Transcription factor G2550 coding sequence, SEQ ID 435.
XX
KW Plant; transcription factor; transgenic plant; abiotic stress tolerance;
KW osmotic stress tolerance; cold tolerance; heat tolerance;
KW low nitrogen tolerance; low phosphate tolerance; fungal disease;
KW glyphosate resistance; flowering; fertility; seed development; ds.
XX
OS Arabidopsis thaliana.
XX
PN WO2004031349-A2.
XX
PD 15-APR-2004.
XX
PF 18-SEP-2003; 2003WO-US030292.
XX
PR 18-SEP-2002; 2002US-0411837P.
PR 17-DEC-2002; 2002US-0434166P.
PR 24-APR-2003; 2003US-0465809P.
XX
PA (MEND-) MENDEL BIOTECHNOLOGY INC.
XX
PI Jiang C, Heard JE, Ratcliffe O, Creelman RA, Adam LJ, Reuber TL;
PI Riechmann JL, Haake V, Dubell AN, Keddie JS, Sherman BK;
XX
DR WPI; 2004-330163/30.
DR P-PSDB; ADO61969.
XX
PT New recombinant polynucleotide encoding transcription factor
PT polypeptides, useful for producing transgenic plants with advantageous
PT properties compared to a reference plant.
XX
PS Claim 1; SEQ ID NO 435; 510pp; English.
XX
CC The present invention relates to novel plant transcription factor
CC proteins (I) and nucleotide sequences (II) (ADO61534-ADO63778). The
CC sequences can be used to produce transgenic plants, which overexpress
CC (II), where the transgenic plant has an altered trait as compared to a
CC non-transgenic plant or wild-type plant. The transgenic plant comprises
CC an altered trait selected from increased tolerance to abiotic stress,
CC increased tolerance to osmotic stress, increased tolerance to cold,
CC increased germination in cold, increased tolerance to heat, increased
CC germination in heat, increased tolerance to freezing conditions,
CC increased tolerance to low nitrogen conditions, increased tolerance to
CC low phosphate conditions, increased tolerance to disease, including
CC fungal disease and particularly Erysiphe, Fusarium and Botrytis,
CC increased tolerance to multiple fungal pathogens, increased resistance to
CC glyphosate, increased sensitivity to ABA, reduced sensitivity to ABA,
CC increased sensitivity to ACC, altered sugar sensing, increased tolerance
CC to sugars, altered carbon/nitrogen sensing, early flowering, late
CC flowering, altered flower structure, loss of flower determinacy, reduced
CC fertility, altered shoot meristem development, altered branching pattern,
CC altered stem morphology, altered vascular tissue structure, reduced
CC apical dominance, altered trichome density, altered trichome development,
CC altered trichome structure, altered root development, altered shade
CC avoidance, altered seed development, altered seed ripening, altered seed
CC germination, slow growth, fast growth, altered cell differentiation,
CC altered cell proliferation, altered cell expansion, altered phase change,
CC altered senescence, abnormal embryo development, altered programmed cell
CC death, lethality when overexpressed, altered necrosis patterns, increased
CC plant size, increased biomass, large seedlings, dwarfed plants, dark
CC green leaves, change in leaf shape, increased leaf size and mass, light
CC green or gray leaves, glossy leaves, altered abaxial/adaxial polarity,
CC altered seed coloration, altered seed size, altered seed shape, large
CC seed, increased leaf wax, increased leaf fatty acids, altered seed oil
CC content, altered seed protein content, altered seedprenyl content,
CC altered leaf prenyl lipid content, increased anthocyanin levels, and
CC decreased anthocyanin levels. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1575 BP; 512 A; 323 C; 323 G; 417 T; 0 U; 0 Other;

Query Match 8.6%; Score 234.4; DB 12; Length 1575;
Best Local Similarity 61.6%; Pred. No. 1.8e-43;

```
  Matches  392;  Conservative    0;  Mismatches  241;  Indels    3;  Gaps    1;

Qy        958 GAGCTTACAACTGCTCAAAGACAAGAACTTCAAATGAAAAAAGCCAAGCTTCTTGCCATG 1017
              ||| |  |  || |  || | || || || || |  || ||   ||||||| |  | |||
Db        670 GAGTTGTCTCCTTCAGAACGTCAGGAGCTACAGAGCAAGAAGAGCAAGCTTTTAACAATG 729

Qy       1018 CTTGAAGAGGTGGAGCAAAGGTACAGACAGTACCATCACCAAATGCAAATAATTGTATTA 1077
              | || ||||| ||  ||||||| |  || |||||||| ||||||  ||    | | ||
Db        730 GTGGATGAGGTAGATAAAAGGTATAACCAATACCATCATCAAATGGAAGCTTTAGCATCG 789

Qy       1078 TCATTTGAGCAAGTAGCAGGAATTGGATCAGCCAAATCATACACTCAATTAGCTTTGCAT 1137
              || || |||   ||| ||||  ||||| |||| ||  | |||||     ||||| || ||
Db        790 TCTTTCGAGATGGTAACAGGTCTTGGAGCAGCTAAGCCTTACACATCCGTAGCTCTGAAT 849

Qy       1138 GCAATTTCGAAGCAATTCAGATGCCTAAAGGATGCAATTGCTGAGCAAGTAAAGGCGACG 1197
              |||  ||    || ||| | ||  ||   || || ||    || ||  |  ||| ||
Db        850 AGAATCTCTCGCCATTTCCGCTGTTTACGCGACGCGATAAAAGAACAGATTCAGGTGATC 909

Qy       1198 AGCAAGAGTTTAGGTGAAGAGGAAGGCTTGGGAGGGAAAATCGAAGG---CTCAAGACTC 1254
              ||   ||   | || ||    ||    |  |  |   ||   ||  |     |  | ||
Db        910 AGAGGGAAGCTTGGGGAGAGAGAGACTTCTGATGAACAAGGAGAGAGGATACCGCGTCTT 969

Qy       1255 AAATTTGTGGACCATCATCTAAGGCAACAACGCGCGCTGCAACAGATAGGAATGATGCAA 1314
              |  |   | || || |   | || |||||  | ||  |||| ||   |       |||
Db        970 AGGTACTTAGATCAACGGTTGAGACAACAGAGAGCTTTGCATCAACAACTTGGAATGGTT 1029

Qy       1315 CCAAATGCTTGGAGACCCCAAAGAGGTTTACCTGAAAGAGCTGTCTCTGTCCTTCGTGCT 1374
              |   |||||||||| |||||||| ||||||||||   |||||||| | ||||| |||
Db       1030 AGACCAGCTTGGAGACCACAAAGAGGCTTACCTGAAAACTCTGTCTCTATACTTCGAGCT 1089

Qy       1375 TGGCTTTTCGAGCATTTTCTTCATCCTTACCCAAAGGATTCAGACAAAATCATGCTTGCT 1434
              ||||| || |||||||| |||||||| || || || || ||||| ||||||||||| |
Db       1090 TGGCTCTTTGAGCATTTCCTTCATCCATATCCTAAAGAATCAGAGAAAATCATGCTTTCA 1149

Qy       1435 AAGCAAACGGGGCTAACAAGGAGCCAGGTGTCTAACTGGTTCATAAATGCTCGAGTTCGA 1494
              ||||| || || ||| | |  | ||||||  | || ||||| || || ||  ||||||||
Db       1150 AAGCAGACAGGACTATCGAAAAACCAGGTTGCAAATTGGTTTATTAACGCGAGAGTTCGA 1209

Qy       1495 TTATGGAAGCCAATGGTAGAAGAAATGTACTTGGAAGAAGTGAAGAATCAAGAACAAAAC 1554
              ||||||| |||||| | ||||| |||||     |||||  |      |   || | |
Db       1210 CTATGGAAACCAATGATTGAAGAGATGTATAAAGAAGAGTTTGGAGAATCAGCAGAGTTA 1269

Qy       1555 AGTACTAATACTTCAGGAGATAACAAAAACAAAGAG 1590
                ||||  ||     ||| | |||||| |   ||
Db       1270 CTCTCTAACTCTAATCAAGACACCAAAAAAATGCAG 1305


RESULT 14
AAC36745
ID   AAC36745 standard; DNA; 1824 BP.
XX
AC   AAC36745;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 14930.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway; metabolic pathway;
KW   promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999;   99US-0121825P.
PR   05-MAR-1999;   99US-0123180P.
PR   09-MAR-1999;   99US-0123548P.
PR   23-MAR-1999;   99US-0125788P.
PR   25-MAR-1999;   99US-0126264P.
PR   29-MAR-1999;   99US-0126785P.
PR   01-APR-1999;   99US-0127462P.
PR   06-APR-1999;   99US-0128234P.
PR   08-APR-1999;   99US-0128714P.
PR   16-APR-1999;   99US-0129845P.
PR   19-APR-1999;   99US-0130077P.
PR   21-APR-1999;   99US-0130449P.
PR   23-APR-1999;   99US-0130510P.
PR   23-APR-1999;   99US-0130891P.
PR   28-APR-1999;   99US-0131449P.
PR   30-APR-1999;   99US-0132048P.
PR   30-APR-1999;   99US-0132407P.
PR   04-MAY-1999;   99US-0132484P.
PR   05-MAY-1999;   99US-0132485P.
PR   06-MAY-1999;   99US-0132486P.
PR   06-MAY-1999;   99US-0132487P.
PR   07-MAY-1999;   99US-0132863P.
PR   11-MAY-1999;   99US-0134256P.
PR   14-MAY-1999;   99US-0134218P.
PR   14-MAY-1999;   99US-0134219P.
PR   14-MAY-1999;   99US-0134221P.
PR   14-MAY-1999;   99US-0134370P.
PR   18-MAY-1999;   99US-0134768P.
PR   19-MAY-1999;   99US-0134941P.
PR   20-MAY-1999;   99US-0135124P.
PR   21-MAY-1999;   99US-0135353P.
PR   24-MAY-1999;   99US-0135629P.
PR   25-MAY-1999;   99US-0136021P.
PR   27-MAY-1999;   99US-0136392P.
PR   28-MAY-1999;   99US-0136782P.
PR   01-JUN-1999;   99US-0137222P.
PR   03-JUN-1999;   99US-0137528P.
PR   04-JUN-1999;   99US-0137502P.
PR   07-JUN-1999;   99US-0137724P.
PR   08-JUN-1999;   99US-0138094P.
PR   10-JUN-1999;   99US-0138540P.
PR   10-JUN-1999;   99US-0138847P.
PR   14-JUN-1999;   99US-0139119P.
PR   16-JUN-1999;   99US-0139452P.
PR   16-JUN-1999;   99US-0139453P.
PR   17-JUN-1999;   99US-0139492P.
PR   18-JUN-1999;   99US-0139454P.
PR   18-JUN-1999;   99US-0139455P.
PR   18-JUN-1999;   99US-0139456P.
PR   18-JUN-1999;   99US-0139457P.
PR   18-JUN-1999;   99US-0139458P.
PR   18-JUN-1999;   99US-0139459P.
PR   18-JUN-1999;   99US-0139460P.
PR   18-JUN-1999;   99US-0139461P.
PR   18-JUN-1999;   99US-0139462P.
PR   18-JUN-1999;   99US-0139463P.
PR   18-JUN-1999;   99US-0139750P.
PR   18-JUN-1999;   99US-0139763P.
PR   21-JUN-1999;   99US-0139817P.
PR   22-JUN-1999;   99US-0139899P.
PR   23-JUN-1999;   99US-0140353P.
PR   23-JUN-1999;   99US-0140354P.
PR   24-JUN-1999;   99US-0140695P.
PR   28-JUN-1999;   99US-0140823P.
PR   29-JUN-1999;   99US-0140991P.
PR   30-JUN-1999;   99US-0141287P.
PR   01-JUL-1999;   99US-0141842P.
PR   01-JUL-1999;   99US-0142154P.
PR   02-JUL-1999;   99US-0142055P.
PR   06-JUL-1999;   99US-0142390P.
PR   08-JUL-1999;   99US-0142803P.
PR   09-JUL-1999;   99US-0142920P.
PR   12-JUL-1999;   99US-0142977P.
PR   13-JUL-1999;   99US-0143542P.
PR   14-JUL-1999;   99US-0143624P.
PR   15-JUL-1999;   99US-0144005P.
PR   16-JUL-1999;   99US-0144085P.
```

PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 8.6%; Score 234.4; DB 3; Length 1824;
Best Local Similarity 61.6%; Pred. No. 1.9e-43;
Matches 392; Conservative 0; Mismatches 241; Indels 3; Gaps 1;



```
Db       1237 CTATGGAAACCAATGATTGAAGAGATGTATAAAGAAGAGTTTGGAGAATCAGCAGAGTTA 1296

Qy       1555 AGTACTAATACTTCAGGAGATAACAAAAACAAAGAG 1590
              |||| ||     ||| | |||||| |   ||
Db       1297 CTCTCTAACTCTAATCAAGACACCAAAAAAATGCAG 1332


RESULT 15
ADO62707
ID   ADO62707 standard; DNA; 2458 BP.
XX
AC   ADO62707;
XX
DT   15-JUL-2004  (first entry)
XX
DE   Transcription factor G2550 orthologous sequence, SEQ ID 1174.
XX
KW   Plant; transcription factor; transgenic plant; abiotic stress tolerance;
KW   osmotic stress tolerance; cold tolerance; heat tolerance;
KW   low nitrogen tolerance; low phosphate tolerance; fungal disease;
KW   glyphosate resistance; flowering; fertility; seed development; ds.
XX
OS   Glycine max.
XX
PN   WO2004031349-A2.
XX
PD   15-APR-2004.
XX
PF   18-SEP-2003; 2003WO-US030292.
XX
PR   18-SEP-2002; 2002US-0411837P.
PR   17-DEC-2002; 2002US-0434166P.
PR   24-APR-2003; 2003US-0465809P.
XX
PA   (MEND-) MENDEL BIOTECHNOLOGY INC.
XX
PI   Jiang C,  Heard JE,  Ratcliffe O,  Creelman RA,  Adam LJ,  Reuber TL;
PI   Riechmann JL,  Haake V,  Dubell AN,  Keddie JS,  Sherman BK;
XX
DR   WPI; 2004-330163/30.
XX
PT   New recombinant polynucleotide encoding transcription factor
PT   polypeptides, useful for producing transgenic plants with advantageous
PT   properties compared to a reference plant.
XX
PS   Claim 1; SEQ ID NO 1174; 510pp; English.
XX
CC   The present invention relates to novel plant transcription factor
CC   proteins (I) and nucleotide sequences (II) (ADO61534-ADO63778). The
CC   sequences can be used to produce transgenic plants, which overexpress
CC   (II), where the transgenic plant has an altered trait as compared to a
CC   non-transgenic plant or wild-type plant. The transgenic plant comprises
CC   an altered trait selected from increased tolerance to abiotic stress,
CC   increased tolerance to osmotic stress, increased tolerance to cold,
CC   increased germination in cold, increased tolerance to heat, increased
CC   germination in heat, increased tolerance to freezing conditions,
CC   increased tolerance to low nitrogen conditions, increased tolerance to
CC   low phosphate conditions, increased tolerance to disease, including
CC   fungal disease and particularly Erysiphe, Fusarium and Botrytis,
CC   increased tolerance to multiple fungal pathogens, increased resistance to
CC   glyphosate, increased sensitivity to ABA, reduced sensitivity to ABA,
CC   increased sensitivity to ACC, altered sugar sensing, increased tolerance
CC   to sugars, altered carbon/nitrogen sensing, early flowering, late
CC   flowering, altered flower structure, loss of flower determinacy, reduced
CC   fertility, altered shoot meristem development, altered branching pattern,
CC   altered stem morphology, altered vascular tissue structure, reduced
CC   apical dominance, altered trichome density, altered trichome development,
CC   altered trichome structure, altered root development, altered shade
CC   avoidance, altered seed development, altered seed ripening, altered seed
CC   germination, slow growth, fast growth, altered cell differentiation,
CC   altered cell proliferation, altered cell expansion, altered phase change,
CC   altered senescence, abnormal embryo development, altered programmed cell
CC   death, lethality when overexpressed, altered necrosis patterns, increased
CC   plant size, increased biomass, large seedlings, dwarfed plants, dark
CC   green leaves, change in leaf shape, increased leaf size and mass, light
CC   green or gray leaves, glossy leaves, altered abaxial/adaxial polarity,
CC   altered seed coloration, altered seed size, altered seed shape, large
CC   seed, increased leaf wax, increased leaf fatty acids, altered seed oil
CC   content, altered seed protein content, altered seedprenyl content,
CC   altered leaf prenyl lipid content, increased anthocyanin levels, and
CC   decreased anthocyanin levels. Note: The sequence data for this patent did
CC   not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 2458 BP; 751 A; 488 C; 545 G; 674 T; 0 U; 0 Other;

  Query Match             8.5%;  Score 231.8;  DB 12;  Length 2458;
  Best Local Similarity   56.8%;  Pred. No. 8.4e-43;
  Matches  481;  Conservative    0;  Mismatches  342;  Indels   24;  Gaps    2;

Qy        741 AAGTTCTAGCAACATGGTTTTAGGCTCTAAGTATCTGAAAGCTGCACAAGAGCTTCTTGA 800
              | |||     ||||  ||||  |||| | || ||||| || ||||||||  | |||||
Db        779 ACGTTACTCAAACAGTATTTTGAACTCTCAATACCTGAAGGCGGCACAAGATTTGCTTGA 838

Qy        801 TGAAGTTGTTAATATTGTTGGAAAAAGCATCAAAGGAGATGATCAAAAGAAGGATAATTC 860
              |||| | | || |                |   ||   |      ||||  | || ||  |
Db        839 TGAAATAGGTAGTGGCCGAAAGGCTTTGAAGCAATCTGGAATGGAAAAACAAGAGAACAC 898

Qy        861 AATGAATAAAGAATCAATGCCTTTGGCTAGTGATGTCAACACTAATAGTTCTGGTGGTGG 920
                  | | || |    ||  | | |  |    | || | |    |
Db        899 TGGATTAGATGGCTCTAAAGATTCTGATGGAAAATCTACCAGTCAATCTATGCAAATGTC 958

Qy        921 TGAAAGTAGCAGCAGGCAGAAAAATGAAGTTGCTGTTGAGCTTACAACTGCTCAAAGACA 980
              | | |    |    |    | | | | ||  ||| ||  || | ||  |  | ||
Db        959 TTCAGGCCCTAATGGTTCCACTGCTAATGCTTCTAGTGAACTATCATCCGCAGAGCGGCA 1018

Qy        981 AGAACTTCAAATGAAAAAAGCCAAGCTTCTTGCCATGCTTGAAGAGGTGGAGCAAAGGTA 1040
              |  |        || ||| | |||||| |  ||||| | || ||||| ||  |||| ||
Db       1019 GAACTTGTTGGACAAGAAAACAAAGCTTTTGTCCATGTTGGATGAGGTAGATAAAAGATA 1078

Qy       1041 CAGACAGTACCATCACCAAATGCAAATAATTGTATTATCATTTGAGCAAGTAGCAGGAAT 1100
              |||||||||   || || ||||| ||  | || | ||| |||||     || || ||
Db       1079 CAGACAGTACTGCCATCAGATGCAGATTGTGGTGTCATCTTTTGACATGGTTGCTGGCTG 1138

Qy       1101 TGGATCAGCCAAATCATACACTCAATTAGCTTTGCATGCAATTTCGAAGCAATTCAGATG 1160
              |||| ||||  || |||| ||  | | || ||     |||||||    || ||  | ||
Db       1139 TGGAGCAGCAGAACCATATACCACACTTGCCTTAAGAACAATTTCTCGCCACTTTCGGTG 1198

Qy       1161 CCTAAAGGATGCAATTGCTGAGCAAGTAAAGGCGACGAGCAAGAGTTTAGGTGAAGAGGA 1220
                |    ||||| ||   ||  ||| |  ||| |||     |  ||  | || ||  |
Db       1199 TTTGCGTGATGCCATCAGTGGCCAAATTCAGGTGACCCAAAGAAGCCTTGGGGAGCA--- 1255

Qy       1221 AGGCTTGGGAGGGAAAATCGAAGGCTCAAGACTCAAATTTGTGGACCATCATCTAAGGCA 1280
                                  | |    |  | |||   | |||||| || || || || ||
Db       1256 ------------------AGGAATACCCCGTCTCCGCTATGTGGATCAGCAACTTAGACA 1297

Qy       1281 ACAACGCGCGCTGCAACAGATAGGAATGATGCAACCAAATGCTTGGAGACCCCAAAGAGG 1340
              ||||   || || || ||  | ||  | ||| ||   | |||||||| || || || ||
Db       1298 ACAAAAGGCACTTCAGCAACTTGGTGTAATGAGAC---AAGCTTGGAGGCCTCAGAGGGG 1354

Qy       1341 TTTACCTGAAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTTTTCGAGCATTTTCTTCATCC 1400
              | | |||||||  |||| ||  | || |||||||||||| || || ||||||||||||||
Db       1355 TCTTCCTGAAAGCTCTGTTTCAATACTCCGTGCTTGGCTCTTTGAACATTTTCTTCATCC 1414

Qy       1401 TTACCCAAAGGATTCAGACAAAATCATGCTTGCTAAGCAAACGGGGCTAACAAGGAGCCA 1460
              |||||| |||||||| || ||||| |||| |||| |||||| ||  | || || | |||
Db       1415 TTACCCCAAGGATTCTGAGAAAATTATGCTGGCTAGGCAAACAGGCTTGACCAGAAACCA 1474

Qy       1461 GGTGTCTAACTGGTTCATAAATGCTCGAGTTCGATTATGGAAGCCAATGGTAGAAGAAAT 1520
              |||| | || |||||||| |||||  | || ||  | ||||||||||||||| || |||||
Db       1475 GGTGGCCAATTGGTTCATCAATGCAAGGGTGCGTCTCTGGAAGCCAATGGTTGAGGAAAT 1534

Qy       1521 GTACTTGGAAGAAGTGAAGAATCAAGAACAAAACAGTACTAATACTTCAGGAGATAACAA 1580
              ||||   || ||  |     |    ||    ||| | |    | | || |   | ||||
Db       1535 GTACAAAGAGGAGTTTGGAGACTCCGAGATGAACTGCAACCTTTCATCTGAGAACAACAC 1594
```

```
Qy       1581 AAACAAA 1587
                 ||||
Db       1595 GGTCAAA 1601
```

Search completed: July 29, 2005, 08:33:46
Job time : 1419 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 29, 2005, 07:15:42 ; Search time 445 Seconds
(without alignments)
10056.667 Million cell updates/sec

Title: US-10-624-201A-1
Perfect score: 2735
Sequence: 1 catgcagagataaaaatata.....gcaaaaaaaaaaaaaaaaaa 2735

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	183.8	6.7	529	4	US-09-640-211A-77	Sequence 77, Appl
	2	164.6	6.0	343	4	US-09-640-211A-1578	Sequence 1578, Ap
	3	159.8	5.8	323	4	US-09-640-211A-418	Sequence 418, App
	4	155.8	5.7	316	4	US-09-640-211A-1854	Sequence 1854, Ap
	5	155.8	5.7	462	4	US-09-640-211A-359	Sequence 359, App
	6	136.2	5.0	1239	4	US-09-640-211A-65	Sequence 65, Appl
	7	134.2	4.9	381	4	US-09-640-211A-102	Sequence 102, App
	8	134.2	4.9	381	4	US-09-640-211A-1472	Sequence 1472, Ap
	9	121	4.4	380	4	US-09-640-211A-111	Sequence 111, App
	10	121	4.4	380	4	US-09-640-211A-1246	Sequence 1246, Ap
	11	120	4.4	690	4	US-09-640-211A-400	Sequence 400, App
	12	111.4	4.1	698	4	US-09-640-211A-37	Sequence 37, Appl
	13	107.4	3.9	366	4	US-09-640-211A-1241	Sequence 1241, Ap
	14	105.4	3.9	474	4	US-09-640-211A-57	Sequence 57, Appl
c	15	105.4	3.9	474	4	US-09-640-211A-1399	Sequence 1399, Ap
	16	96	3.5	171	4	US-09-640-211A-1845	Sequence 1845, Ap
	17	89.4	3.3	158	4	US-09-640-211A-370	Sequence 370, App
	18	84.6	3.1	260	4	US-09-640-211A-351	Sequence 351, App
	19	84.6	3.1	260	4	US-09-640-211A-1833	Sequence 1833, Ap
	20	84.2	3.1	157	4	US-09-640-211A-394	Sequence 394, App
	21	81.6	3.0	486	4	US-09-640-211A-50	Sequence 50, Appl
	22	60.6	2.2	533	4	US-09-640-211A-185	Sequence 185, App
	23	59.8	2.2	30820	4	US-09-949-016-17145	Sequence 17145, A
	24	59.4	2.2	407	4	US-09-640-211A-383	Sequence 383, App
c	25	59.2	2.2	7218	1	US-08-232-463-14	Sequence 14, Appl
c	26	58.2	2.1	18773	4	US-09-949-016-14164	Sequence 14164, A
c	27	56.4	2.1	19124	2	US-08-487-826B-13	Sequence 13, Appl
c	28	56	2.0	612	4	US-09-902-540-1357	Sequence 1357, Ap
	29	55.6	2.0	3190	4	US-09-949-016-5219	Sequence 5219, Ap
	30	55.6	2.0	3439	4	US-09-949-016-648	Sequence 648, App
c	31	55	2.0	187169	4	US-09-949-016-12776	Sequence 12776, A
c	32	55	2.0	191569	4	US-09-949-016-15940	Sequence 15940, A
	33	54	2.0	2905	4	US-09-976-594-1031	Sequence 1031, Ap
	34	54	2.0	2905	4	US-09-919-039-381	Sequence 381, App
c	35	53.8	2.0	30820	4	US-09-949-016-17145	Sequence 17145, A
c	36	53.2	1.9	1141	4	US-09-806-708B-22	Sequence 22, Appl
c	37	53.2	1.9	67755	4	US-09-949-016-13703	Sequence 13703, A
	38	53	1.9	273	4	US-09-313-294A-2156	Sequence 2156, Ap
	39	52.8	1.9	119153	4	US-09-949-016-12378	Sequence 12378, A
	40	51.2	1.9	1141	4	US-09-806-708B-22	Sequence 22, Appl
	41	51	1.9	1770	4	US-09-976-594-195	Sequence 195, App
	42	50.8	1.9	19438	4	US-09-949-016-12699	Sequence 12699, A
c	43	50.2	1.8	601	4	US-09-949-016-168053	Sequence 168053,
c	44	50.2	1.8	601	4	US-09-949-016-168054	Sequence 168054,
c	45	50.2	1.8	601	4	US-09-949-016-168055	Sequence 168055,

ALIGNMENTS

RESULT 1
US-09-640-211A-77
; Sequence 77, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; TITLE OF INVENTION: Modification of Gene Transcription
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 77
; LENGTH: 529
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-77

Query Match 6.7%; Score 183.8; DB 4; Length 529;
Best Local Similarity 77.0%; Pred. No. 2.6e-37;
Matches 224; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

```
Qy       1258 TTTGTGGACCATCATCTAAGGCAACAACGCGCGCTGCAACAGATAGGAATGATGCAACCA 1317
             ||| ||||||| || || ||||| ||||| || || || ||| | || ||||||||||
Db         38 TTTTTGGACCAGCAGCTCAGGCAGCAACGAGCCCTACAGCAGCTTGGGATGATGCAACAG 97

Qy       1318 AATGCTTGGAGACCCCAAAGAGGTTTACCTGAAAGAGCTGTCTCTGTCCTTCGTGCTTGG 1377
              |||| ||||| || |||||||| | ||||| || |||| ||| | ||||| || |||
Db         98 CATGCATGGAGGCCGCAAAGAGGACTTCCTGAGAGTTCTGTTTCTATTCTTCGGGCCTGG 157

Qy       1378 CTTTTCGAGCATTTTCTTCATCCTTACCCAAAGGATTCAGACAAAATCATGCTTGCTAAG 1437
              || || ||||||||||||||||||||||||||||||| ||||||||| ||||||| | |
Db        158 CTATTTGAGCATTTTCTTCATCCTTACCCAAAGGATTCTGACAAAATCCTGCTTGCAAGG 217

Qy       1438 CAAACGGGGCTAACAAGGAGCCAGGTGTCTAACTGGTTCATAAATGCTCGAGTTCGATTA 1497
              || || ||  | ||||| || ||||| || || |||||||| |||||  |||| ||  |
Db        218 CAGACAGGCTTGACAAGAAGTCAGGTCTCGAATTGGTTCATCAATGCAAGAGTGCGTCTC 277

Qy       1498 TGGAAGCCAATGGTAGAAGAAATGTACTTGGAAGAAGTGAAGAATCAAGAA 1548
              ||||| || ||||| ||||||||||||   |||||  |   | ||   |||
Db        278 TGGAAACCTATGGTCGAAGAAATGTACAAAGAAGAGATTGGGGATGCGGAA 328
```

RESULT 2
US-09-640-211A-1578
; Sequence 1578, Application US/09640211A

```
; Patent No. 6833446
; GENERAL INFORMATION:
;  APPLICANT: Wood, Marion
;  APPLICANT:  Shenk, Michael A.
;  APPLICANT:  McGrath, Annette
;  APPLICANT:  Glenn, Matthew
;  TITLE OF INVENTION: Compositions and Methods for the
;  TITLE OF INVENTION:  Modification of Gene Transcription
;  FILE REFERENCE: 11000.1021C1U
;  CURRENT APPLICATION NUMBER: US/09/640,211A
;  CURRENT FILING DATE:  2000-08-16
;  NUMBER OF SEQ ID NOS: 2368
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1578
;   LENGTH: 343
;   TYPE: DNA
;   ORGANISM: Pinus radiata
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(343)
;   OTHER INFORMATION: n = A,T,C or G
US-09-640-211A-1578

  Query Match             6.0%;  Score 164.6;  DB 4;  Length 343;
  Best Local Similarity   70.6%;  Pred. No. 1.9e-32;
  Matches  218;  Conservative    0;  Mismatches   91;  Indels    0;  Gaps    0;

Qy       1240 GAAGGCTCAAGACTCAAATTTGTGGACCATCATCTAAGGCAACAACGCGCGCTGCAACAG 1299
              |||    ||||  |    || | | || |       || | ||||| |||||  | || ||
Db          1 GAAACACCAAGGTTGGGATNTCTAGAACGAAGCATACGACAACAGCGCGCATTTCACCAC 60

Qy       1300 ATAGGAATGATGCAACCAAATGCTTGGAGACCCCAAAGAGGTTTACCTGAAAGAGCTGTC 1359
              |||||  |||||  | |    |  |||||  |||| || |||||  | ||||||  |  ||||
Db         61 TTAGGATTGATGGAGCAGCACCCTTGGCGACCGCAGAGAGGACTTCCTGAACGCTCTGTN 120

Qy       1360 TCTGTCCTTCGTGCTTGGCTTTTCGAGCATTTTCTTCATCCTTACCCAAAGGATTCAGAC 1419
              |||||  |||||||| |||  | || ||||||||||| || || || ||||  ||| ||||
Db        121 TCTGTTCTTCGTGCATGGTTGTTTGAGCATTTTCTGCACCCGTATCCAACTGATGCAGAT 180

Qy       1420 AAAATCATGCTTGCTAAGCAAACGGGGCTAACAAGGAGCCAGGTGTCTAACTGGTTCATA 1479
              ||    ||  | ||||||||||| || || |||||  || |||||  || || |||||  |||
Db        181 AAGCATATATTGGCTAAGCAAACTGGCCTTACAAGAAGTCAGGTATCAAATTGGTTTATA 240

Qy       1480 AATGCTCGAGTTCGATTATGGAAGCCAATGGTAGAAGAAATGTACTTGGAAGAAGTGAAG 1539
              |||||  | |||  || |||||||||| |||||  || || ||||||  |||||||| | |||
Db        241 AATGCCAGGGTTAGACTATGGAAGCCCATGGTGGAGGAGATGTACATGGAAGAACTCAAG 300

Qy       1540 AATCAAGAA 1548
                |  || ||
Db        301 GAAGAAAAA 309


RESULT 3
US-09-640-211A-418
; Sequence 418, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
;  APPLICANT: Wood, Marion
;  APPLICANT:  Shenk, Michael A.
;  APPLICANT:  McGrath, Annette
;  APPLICANT:  Glenn, Matthew
;  TITLE OF INVENTION: Compositions and Methods for the
;  TITLE OF INVENTION:  Modification of Gene Transcription
;  FILE REFERENCE: 11000.1021C1U
;  CURRENT APPLICATION NUMBER: US/09/640,211A
;  CURRENT FILING DATE:  2000-08-16
;  NUMBER OF SEQ ID NOS: 2368
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 418
;   LENGTH: 323
;   TYPE: DNA
;   ORGANISM: Pinus radiata
US-09-640-211A-418

  Query Match             5.8%;  Score 159.8;  DB 4;  Length 323;
  Best Local Similarity   73.8%;  Pred. No. 3.1e-31;
  Matches  203;  Conservative    0;  Mismatches   72;  Indels    0;  Gaps    0;

Qy       1274 TAAGGCAACAACGCGCGCTGCAACAGATAGGAATGATGCAACCAAATGCTTGGAGACCCC 1333
              || | ||||| |||||  | || ||  ||||| |||||  | |    |  ||||| |||| |
Db         15 TACGACAACAGCGCGCATTTCACCACTTAGGATTGATGGAGCAGCACCCTTGGCGACCGC 74

Qy       1334 AAAGAGGTTTACCTGAAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTTTTCGAGCATTTTC 1393
              | |||||  | ||||||  |  |||| |||||  |||||||| |||  | || ||||||||||
Db         75 AGAGAGGACTTCCTGAACGCTCTGTTTCTGTTCTTCGTGCATGGTTGTTTGAGCATTTTC 134

Qy       1394 TTCATCCTTACCCAAAGGATTCAGACAAAATCATGCTTGCTAAGCAAACGGGGCTAACAA 1453
              | || || || ||||  ||| |||| ||    ||  | ||||||||||| || || ||||
Db        135 TGCACCCGTATCCAACTGATGCAGATAAGCATATATTGGCTAAGCAAACTGGCCTTACAA 194

Qy       1454 GGAGCCAGGTGTCTAACTGGTTCATAAATGCTCGAGTTCGATTATGGAAGCCAATGGTAG 1513
              | || ||||| || || ||||| ||||||||  | ||| || ||||||||||| |||||  |
Db        195 GAAGTCAGGTATCAAATTGGTTTATAAATGCCAGGGTTAGACTATGGAAGCCCATGGTGG 254

Qy       1514 AAGAAATGTACTTGGAAGAAGTGAAGAATCAAGAA 1548
              | || |||||| |||||||| | ||| |  || ||
Db        255 AGGAGATGTACATGGAAGAACTCAAGGAAGAAAAA 289


RESULT 4
US-09-640-211A-1854
; Sequence 1854, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
;  APPLICANT: Wood, Marion
;  APPLICANT:  Shenk, Michael A.
;  APPLICANT:  McGrath, Annette
;  APPLICANT:  Glenn, Matthew
;  TITLE OF INVENTION: Compositions and Methods for the
;  TITLE OF INVENTION:  Modification of Gene Transcription
;  FILE REFERENCE: 11000.1021C1U
;  CURRENT APPLICATION NUMBER: US/09/640,211A
;  CURRENT FILING DATE:  2000-08-16
;  NUMBER OF SEQ ID NOS: 2368
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1854
;   LENGTH: 316
;   TYPE: DNA
;   ORGANISM: Pinus radiata
US-09-640-211A-1854

  Query Match             5.7%;  Score 155.8;  DB 4;  Length 316;
  Best Local Similarity   73.4%;  Pred. No. 3.3e-30;
  Matches  199;  Conservative    0;  Mismatches   72;  Indels    0;  Gaps    0;

Qy       1284 ACGCGCGCTGCAACAGATAGGAATGATGCAACCAAATGCTTGGAGACCCCAAAGAGGTTT 1343
              ||| || || |||||  |||| ||||| || |   |||||||||| || || |||||  |
Db          1 ACGGGCTCTCCAACAATTAGGCATGATTCAGCAGCATGCTTGGAGGCCACAGAGAGGACT 60

Qy       1344 ACCTGAAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTTTTCGAGCATTTTCTTCATCCTTA 1403
              || ||  || |||| |||||| | || ||||||||  || || ||||||||||||||  ||
Db         61 TCCCGAGCGATCTGTTTCTGTCTTACGGGCTTGGCTATTTGAACATTTTCTTCATCCGTA 120

Qy       1404 CCCAAAGGATTCAGACAAAATCATGCTTGCTAAGCAAACGGGGCTAACAAGGAGCCAGGT 1463
              |||||  |||  |||||||||    ||||| || |  || || |||||  || || |  |||||
Db        121 TCCAAAAGATGCAGACAAACATATGCTCGCGAGACAGACTGGGCTTACCAGAAATCAGGT 180

Qy       1464 GTCTAACTGGTTCATAAATGCTCGAGTTCGATTATGGAAGCCAATGGTAGAAGAAATGTA 1523
              || || ||||| |||||||| || || ||  | |||||||| ||||| ||||| |||||
Db        181 CTCAAATTGGTTTATAAATGCACGTGTACGCCTCTGGAAGCCTATGGTGGAAGAGATGTA 240

Qy       1524 CTTGGAAGAAGTGAAGAATCAAGAACAAAAC 1554
              |||| |||   ||| |  ||||  | ||
Db        241 TGTGGAGGAAACAAAGGAGGCAGAAGTAGAC 271
```

RESULT 5
US-09-640-211A-359
; Sequence 359, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; TITLE OF INVENTION: Modification of Gene Transcription
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 359
; LENGTH: 462
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-640-211A-359

Query Match 5.7%; Score 155.8; DB 4; Length 462;
Best Local Similarity 73.4%; Pred. No. 4.1e-30;
Matches 199; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

```
Qy       1284 ACGCGCGCTGCAACAGATAGGAATGATGCAACCAAATGCTTGGAGACCCCAAAGAGGTTT 1343
Db          1 ACGGGCTCTCCAACAATTAGGCATGATTCAGCAGCATGCTTGGAGGCCACAGAGAGGACT 60

Qy       1344 ACCTGAAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTTTTCGAGCATTTTCTTCATCCTTA 1403
Db         61 TCCCGAGCGATCTGTTTCTGTCTTACGGGCTTGGCTATTTGAACATTTTCTTCATCCGTA 120

Qy       1404 CCCAAAGGATTCAGACAAAATCATGCTTGCTAAGCAAACGGGGCTAACAAGGAGCCAGGT 1463
Db        121 TCCAAAAGATGCAGACAAACATATGCTCGCGAGACAGACTGGGCTTACCAGAAATCAGGT 180

Qy       1464 GTCTAACTGGTTCATAAATGCTCGAGTTCGATTATGGAAGCCAATGGTAGAAGAAATGTA 1523
Db        181 CTCAAATTGGTTTATAAATGCACGTGTACGCCTCTGGAAGCCTATGGTGGAAGAGATGTA 240

Qy       1524 CTTGGAAGAAGTGAAGAATCAAGAACAAAAC 1554
Db        241 TGTGGAGGAAACAAAGGAGGCAGAAGTAGAC 271
```

RESULT 6
US-09-640-211A-65
; Sequence 65, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; TITLE OF INVENTION: Modification of Gene Transcription
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 65
; LENGTH: 1239
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-65

Query Match 5.0%; Score 136.2; DB 4; Length 1239;
Best Local Similarity 59.4%; Pred. No. 7.6e-25;
Matches 280; Conservative 0; Mismatches 168; Indels 23; Gaps 2;

```
Qy        755 TGGTTTTAGGCTCTAAGTATCTGAAAGCTGCACAAGAGCTTCTTGATGAAGTTGTTAATA 814
Db        790 TGTTGCTGAGCTCCAAGTACTTGAAGGCCGCACAGGAGCTTCTCGAAGAGGTTGTTAATG 849

Qy        815 TTGTTGGAAAAAGCATCAAAGGAGATGATCAAAAGAAGGATAATTCAATGAATAAAGAAT 874
Db        850 TTGGGAACACCGGAATCAAAGCTGAGATGCTGAAGAAGGCCAGCGGCCAAAGTAA----- 904

Qy        875 CAATGCCTTTGGCTAGTGATGTCAACACTAATAGTTCTGGTGGTGGTGAAAGTAGCAGCA 934
Db        905 -----------------GCCGGGTGGAGAATCAGCGGCACTGAAGGAGGAAGGAGGTGGC 947

Qy        935 GGCAGAAAAATGAAG-TTGCTGTTGAGCTTACAACTGCTCAAAGACAAGAACTTCAAATG 993
Db        948 GACGGCAGTGGTAAGCGCGGCGCAGAACTGTCTATGGCTGAGAGGCAGGAAATTCAGATG 1007

Qy        994 AAAAAAGCCAAGCTTCTTGCCATGCTTGAAGAGGTGGAGCAAAGGTACAGACAGTACCAT 1053
Db       1008 AAGAAGGCTAAGCTCATAAACATGCTTGATGAGGTTGAACAGAGATACAGACAGTACCAC 1067

Qy       1054 CACCAAATGCAAATAATTGTATTATCATTTGAGCAAGTAGCAGGAATTGGATCAGCCAAA 1113
Db       1068 AACCAGATGCAGATCGTGATATCCTCGTTTGAGCAAGCTGCGGGGATTGGCTCTGCAAGG 1127

Qy       1114 TCATACACTCAATTAGCTTTGCATGCAATTTCGAAGCAATTCAGATGCCTAAAGGATGCA 1173
Db       1128 ACGTACACGGCCCTCGCGCTGCAGACCATCTCAAAGCAGTTCCGGTGCCTCAAGGACGCG 1187

Qy       1174 ATTGCTGAGCAAGTAAAGGCGACGAGCAAGAGTTTAGGTGAAGAGGAAGGC 1224
Db       1188 ATAGCAGGCCAGATTCGGGCCGCTAATAAGAGCTTGGGCGAGGAGGATGGC 1238
```

RESULT 7
US-09-640-211A-102
; Sequence 102, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; TITLE OF INVENTION: Modification of Gene Transcription
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 102
; LENGTH: 381
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-102

Query Match 4.9%; Score 134.2; DB 4; Length 381;
Best Local Similarity 65.6%; Pred. No. 1.4e-24;
Matches 196; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

```
Qy       1234 AAAATCGAAGGCTCAAGACTCAAATTTGTGGACCATCATCTAAGGCAACAACGCGCGCTG 1293
Db         72 AAGGGCGAGACACCACGGCTCAAGTTGCTCGACCAGAGCCTGAGGCAGCAGAGGGCTTTC 131

Qy       1294 CAACAGATAGGAATGATGCAACCAAATGCTTGGAGACCCCAAAGAGGTTTACCTGAAAGA 1353
Db        132 CACCAGATGGGCATGATGGAGCAAGAGGCCTGGAGGCCGCAGCGGGGCCTGCCGGAGCGG 191

Qy       1354 GCTGTCTCTGTCCTTCGTGCTTGGCTTTTCGAGCATTTTCTTCATCCTTACCCAAAGGAT 1413
Db        192 TCGGTCAACATACTGCGTGCATGGCTCTTCGAGCATTTCTTGCATCCGTATCCAAGTGAC 251

Qy       1414 TCAGACAAAATCATGCTTGCTAAGCAAACGGGGCTAACAAGGAGCCAGGTGTCTAACTGG 1473
```

```
Db        252 GCTGATAAGCATCTGTTGGCTCGACAGACTGGTCTCTCCAGAAACCAGGTCTCGAATTGG 311

Qy       1474 TTCATAAATGCTCGAGTTCGATTATGGAAGCCAATGGTAGAAGAAATGTACTTGGAAGA 1532
              ||||||||||||  | || || || ||||| || ||||| || || ||||||  | ||||
Db        312 TTCATAAATGCCAGGGTCCGGTTGTGGAAACCCATGGTGGAGGAGATGTACCAGCAAGA 370


RESULT 8
US-09-640-211A-1472
; Sequence 1472, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
;  APPLICANT: Wood, Marion
;  APPLICANT:  Shenk, Michael A.
;  APPLICANT:  McGrath, Annette
;  APPLICANT:  Glenn, Matthew
;  TITLE OF INVENTION: Compositions and Methods for the
;  TITLE OF INVENTION:  Modification of Gene Transcription
;  FILE REFERENCE: 11000.1021C1U
;  CURRENT APPLICATION NUMBER: US/09/640,211A
;  CURRENT FILING DATE:  2000-08-16
;  NUMBER OF SEQ ID NOS: 2368
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1472
;   LENGTH: 381
;   TYPE: DNA
;   ORGANISM: Eucalyptus grandis
US-09-640-211A-1472

  Query Match             4.9%;  Score 134.2;  DB 4;  Length 381;
  Best Local Similarity   65.6%;  Pred. No. 1.4e-24;
  Matches  196;  Conservative    0;  Mismatches  103;  Indels    0;  Gaps    0;

Qy       1234 AAAATCGAAGGCTCAAGACTCAAATTTGTGGACCATCATCTAAGGCAACAACGCGCGCTG 1293
              ||   |||     || | ||||| ||  | |||||    || ||||| ||  | ||  |
Db         72 AAGGGCGAGACACCACGGCTCAAGTTGCTCGACCAGAGCCTGAGGCAGCAGAGGGCTTTC 131

Qy       1294 CAACAGATAGGAATGATGCAACCAAATGCTTGGAGACCCCAAAGAGGTTTACCTGAAAGA 1353
              || ||||| || |||||| | | | | || ||||| || ||  | ||  | || ||  |
Db        132 CACCAGATGGGCATGATGGAGCAAGAGGCCTGGAGGCCGCAGCGGGGCCTGCCGGAGCGG 191

Qy       1354 GCTGTCTCTGTCCTTCGTGCTTGGCTTTTCGAGCATTTTCTTCATCCTTACCCAAAGGAT 1413
              | |||    | || ||||| ||||| ||||||||||  | ||||| || ||||  ||
Db        192 TCGGTCAACATACTGCGTGCATGGCTCTTCGAGCATTTCTTGCATCCGTATCCAAGTGAC 251

Qy       1414 TCAGACAAAATCATGCTTGCTAAGCAAACGGGGCTAACAAGGAGCCAGGTGTCTAACTGG 1473
              | || ||     || | |||   || || || ||  | || | |||||| || || |||
Db        252 GCTGATAAGCATCTGTTGGCTCGACAGACTGGTCTCTCCAGAAACCAGGTCTCGAATTGG 311

Qy       1474 TTCATAAATGCTCGAGTTCGATTATGGAAGCCAATGGTAGAAGAAATGTACTTGGAAGA 1532
              ||||||||||||  | || || || ||||| || ||||| || || ||||||  | ||||
Db        312 TTCATAAATGCCAGGGTCCGGTTGTGGAAACCCATGGTGGAGGAGATGTACCAGCAAGA 370


RESULT 9
US-09-640-211A-111
; Sequence 111, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
;  APPLICANT: Wood, Marion
;  APPLICANT:  Shenk, Michael A.
;  APPLICANT:  McGrath, Annette
;  APPLICANT:  Glenn, Matthew
;  TITLE OF INVENTION: Compositions and Methods for the
;  TITLE OF INVENTION:  Modification of Gene Transcription
;  FILE REFERENCE: 11000.1021C1U
;  CURRENT APPLICATION NUMBER: US/09/640,211A
;  CURRENT FILING DATE:  2000-08-16
;  NUMBER OF SEQ ID NOS: 2368
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 111
;   LENGTH: 380
;   TYPE: DNA
;   ORGANISM: Eucalyptus grandis
US-09-640-211A-111

  Query Match             4.4%;  Score 121;  DB 4;  Length 380;
  Best Local Similarity   80.2%;  Pred. No. 3.4e-21;
  Matches  142;  Conservative    0;  Mismatches   35;  Indels    0;  Gaps    0;

Qy       1377 GCTTTTCGAGCATTTTCTTCATCCTTACCCAAAGGATTCAGACAAAATCATGCTTGCTAA 1436
              ||| ||||| || ||||| || || ||||| |||||||| |||||| ||||||| || ||
Db          1 GCTCTTCGAACACTTTCTCCACCCCTACCCGAAGGATTCGGACAAAGTCATGCTGGCCAA 60

Qy       1437 GCAAACGGGGCTAACAAGGAGCCAGGTGTCTAACTGGTTCATAAATGCTCGAGTTCGATT 1496
              || || ||||| || || |||||||||||| || ||||| ||||||||||||||||||  |
Db         61 ACAGACAGGGCTCACTAGAAGCCAGGTGTCGAATTGGTTTATAAATGCTCGAGTTCGGCT 120

Qy       1497 ATGGAAGCCAATGGTAGAAGAAATGTACTTGGAAGAAGTGAAGAATCAAGAACAAAA 1553
              |||||||| ||||| || || ||||||  ||| ||| | ||| | |||||||| ||
Db        121 TTGGAAGCCGATGGTGGAGGAGATGTACACGGAGGAAATCAAGGAGCAAGAACAGAA 177


RESULT 10
US-09-640-211A-1246
; Sequence 1246, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
;  APPLICANT: Wood, Marion
;  APPLICANT:  Shenk, Michael A.
;  APPLICANT:  McGrath, Annette
;  APPLICANT:  Glenn, Matthew
;  TITLE OF INVENTION: Compositions and Methods for the
;  TITLE OF INVENTION:  Modification of Gene Transcription
;  FILE REFERENCE: 11000.1021C1U
;  CURRENT APPLICATION NUMBER: US/09/640,211A
;  CURRENT FILING DATE:  2000-08-16
;  NUMBER OF SEQ ID NOS: 2368
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1246
;   LENGTH: 380
;   TYPE: DNA
;   ORGANISM: Eucalyptus grandis
US-09-640-211A-1246

  Query Match             4.4%;  Score 121;  DB 4;  Length 380;
  Best Local Similarity   80.2%;  Pred. No. 3.4e-21;
  Matches  142;  Conservative    0;  Mismatches   35;  Indels    0;  Gaps    0;

Qy       1377 GCTTTTCGAGCATTTTCTTCATCCTTACCCAAAGGATTCAGACAAAATCATGCTTGCTAA 1436
              ||| ||||| || ||||| || || ||||| |||||||| |||||| ||||||| || ||
Db          1 GCTCTTCGAACACTTTCTCCACCCCTACCCGAAGGATTCGGACAAAGTCATGCTGGCCAA 60

Qy       1437 GCAAACGGGGCTAACAAGGAGCCAGGTGTCTAACTGGTTCATAAATGCTCGAGTTCGATT 1496
              || || ||||| || || |||||||||||| || ||||| ||||||||||||||||||  |
Db         61 ACAGACAGGGCTCACTAGAAGCCAGGTGTCGAATTGGTTTATAAATGCTCGAGTTCGGCT 120

Qy       1497 ATGGAAGCCAATGGTAGAAGAAATGTACTTGGAAGAAGTGAAGAATCAAGAACAAAA 1553
              |||||||| ||||| || || ||||||  ||| ||| | ||| | |||||||| ||
Db        121 TTGGAAGCCGATGGTGGAGGAGATGTACACGGAGGAAATCAAGGAGCAAGAACAGAA 177


RESULT 11
US-09-640-211A-400
; Sequence 400, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
;  APPLICANT: Wood, Marion
;  APPLICANT:  Shenk, Michael A.
;  APPLICANT:  McGrath, Annette
;  APPLICANT:  Glenn, Matthew
;  TITLE OF INVENTION: Compositions and Methods for the
;  TITLE OF INVENTION:  Modification of Gene Transcription
;  FILE REFERENCE: 11000.1021C1U
;  CURRENT APPLICATION NUMBER: US/09/640,211A
;  CURRENT FILING DATE:  2000-08-16
```

```
;  NUMBER OF SEQ ID NOS: 2368
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 400
;   LENGTH: 690
;   TYPE: DNA
;   ORGANISM: Pinus radiata
US-09-640-211A-400

  Query Match             4.4%;  Score 120;  DB 4;  Length 690;
  Best Local Similarity   78.3%;  Pred. No. 8.4e-21;
  Matches  144;  Conservative    0;  Mismatches   40;  Indels    0;  Gaps    0;

Qy        939 GAAAAATGAAGTTGCTGTTGAGCTTACAACTGCTCAAAGACAAGAACTTCAAATGAAAAA 998
              |||||| ||    |||| |||  || ||| | ||| | ||||||||||||||| ||||||||
Db        504 GAAAAACGAGCATGCTTTTGGACTCACACCAGCTGATAGACAAGAACTTCAGATGAAAAA 563

Qy        999 AGCCAAGCTTCTTGCCATGCTTGAAGAGGTGGAGCAAAGGTACAGACAGTACCATCACCA 1058
              || |||||| |||||||| | || |||||||| | |||||||||||||||| |||| ||
Db        564 GGCAAAGCTTGTTGCCATGTTGGATGAGGTGGATCGAAGGTACAGACAGTACTATCATCA 623

Qy       1059 AATGCAAATAATTGTATTATCATTTGAGCAAGTAGCAGGAATTGGATCAGCCAAATCATA 1118
              |||||||| |||| | ||| ||||||    | ||| ||| |||| | |||||  ||||
Db        624 GATGCAAATCGTTGTTTCATCGTTTGAGACCGCAGCTGGATTTGGGGCTGCCAAGACATA 683

Qy       1119 CACT 1122
              ||||
Db        684 CACT 687


RESULT 12
US-09-640-211A-37
; Sequence 37, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
;  APPLICANT: Wood, Marion
;  APPLICANT:  Shenk, Michael A.
;  APPLICANT:  McGrath, Annette
;  APPLICANT:  Glenn, Matthew
;  TITLE OF INVENTION: Compositions and Methods for the
;  TITLE OF INVENTION:  Modification of Gene Transcription
;  FILE REFERENCE: 11000.1021C1U
;  CURRENT APPLICATION NUMBER: US/09/640,211A
;  CURRENT FILING DATE:  2000-08-16
;  NUMBER OF SEQ ID NOS: 2368
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 37
;   LENGTH: 698
;   TYPE: DNA
;   ORGANISM: Eucalyptus grandis
US-09-640-211A-37

  Query Match             4.1%;  Score 111.4;  DB 4;  Length 698;
  Best Local Similarity   63.0%;  Pred. No. 1.4e-18;
  Matches  172;  Conservative    0;  Mismatches  101;  Indels    0;  Gaps    0;

Qy        953 CTGTTGAGCTTACAACTGCTCAAAGACAAGAACTTCAAATGAAAAAAGCCAAGCTTCTTG 1012
              ||   |||||| |  ||||| ||| ||||||  | || |  ||     | |||||  |
Db        415 CTAGCGAGCTTTCTGCTGCTGAAAAACAAGATTTGCAGAACAAGCTCACGAAGCTCTTGT 474

Qy       1013 CCATGCTTGAAGAGGTGGAGCAAAGGTACAGACAGTACCATCACCAAATGCAAATAATTG 1072
              ||||| | || ||||| ||  |||||||||  |||||| |||||||||||||  ||  | |
Db        475 CCATGTTGGACGAGGTTGATAAAAGGTACAAGCAGTACTATCACCAAATGCAGATCGTGG 534

Qy       1073 TATTATCATTTGAGCAAGTAGCAGGAATTGGATCAGCCAAATCATACACTCAATTAGCTT 1132
              ||   || |||||   | |||||||||  ||  |||||||  | |||||     | ||
Db        535 TACAGTCTTTTGATACAATAGCAGGAAGCGGTGCAGCCAAGCCCTACACGGCGCTTGCGC 594

Qy       1133 TGCATGCAATTTCGAAGCAATTCAGATGCCTAAAGGATGCAATTGCTGAGCAAGTAAAGG 1192
              | ||    || ||    || ||| | |||||  | || || ||  | |  ||| |  | |
Db        595 TCCAGAGGATATCCCGCCACTTCCGGTGCCTGCATGACGCGATCACGGGTCAAATTCAAG 654

Qy       1193 CGACGAGCAAGAGTTTAGGTGAAGAGGAAGGCT 1225
              | ||  | || ||| | || ||  | ||   ||
Db        655 CAACCCGTAAAAGTCTCGGAGAGCAAGACACCT 687


RESULT 13
US-09-640-211A-1241
; Sequence 1241, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
;  APPLICANT: Wood, Marion
;  APPLICANT:  Shenk, Michael A.
;  APPLICANT:  McGrath, Annette
;  APPLICANT:  Glenn, Matthew
;  TITLE OF INVENTION: Compositions and Methods for the
;  TITLE OF INVENTION:  Modification of Gene Transcription
;  FILE REFERENCE: 11000.1021C1U
;  CURRENT APPLICATION NUMBER: US/09/640,211A
;  CURRENT FILING DATE:  2000-08-16
;  NUMBER OF SEQ ID NOS: 2368
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1241
;   LENGTH: 366
;   TYPE: DNA
;   ORGANISM: Eucalyptus grandis
US-09-640-211A-1241

  Query Match             3.9%;  Score 107.4;  DB 4;  Length 366;
  Best Local Similarity   78.2%;  Pred. No. 1.1e-17;
  Matches  129;  Conservative    0;  Mismatches   36;  Indels    0;  Gaps    0;

Qy       1384 GAGCATTTTCTTCATCCTTACCCAAAGGATTCAGACAAAATCATGCTTGCTAAGCAAACG 1443
              |||||||||||||||||||||||||||||||||||||||| |||||||| ||||||| | ||| ||
Db          1 GAGCATTTTCTTCATCCTTACCCAAAGGATTCTGACAAAATCCTGCTTGCAAGGCAGACA 60

Qy       1444 GGGCTAACAAGGAGCCAGGTGTCTAACTGGTTCATAAATGCTCGAGTTCGATTATGGAAG 1503
              ||  | ||||| || ||||| || || |||||||| |||||  |||| ||  | |||||
Db         61 GGCTTGACAAGAAGTCAGGTCTCGAATTGGTTCATCAATGCAAGAGTGCGTCTCTGGAAA 120

Qy       1504 CCAATGGTAGAAGAAATGTACTTGGAAGAAGTGAAGAATCAAGAA 1548
              || ||||| ||||||||||||   |||||  |   | ||   |||
Db         121 CCTATGGTCGAAGAAATGTACAAAGAAGAGATTGGGGATGCGGAA 165


RESULT 14
US-09-640-211A-57
; Sequence 57, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
;  APPLICANT: Wood, Marion
;  APPLICANT:  Shenk, Michael A.
;  APPLICANT:  McGrath, Annette
;  APPLICANT:  Glenn, Matthew
;  TITLE OF INVENTION: Compositions and Methods for the
;  TITLE OF INVENTION:  Modification of Gene Transcription
;  FILE REFERENCE: 11000.1021C1U
;  CURRENT APPLICATION NUMBER: US/09/640,211A
;  CURRENT FILING DATE:  2000-08-16
;  NUMBER OF SEQ ID NOS: 2368
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 57
;   LENGTH: 474
;   TYPE: DNA
;   ORGANISM: Eucalyptus grandis
US-09-640-211A-57

  Query Match             3.9%;  Score 105.4;  DB 4;  Length 474;
  Best Local Similarity   77.9%;  Pred. No. 4e-17;
  Matches  127;  Conservative    0;  Mismatches   36;  Indels    0;  Gaps    0;

Qy       1402 TACCCAAAGGATTCAGACAAAATCATGCTTGCTAAGCAAACGGGGCTAACAAGGAGCCAG 1461
              || || || ||||| ||||||  |||||| || || ||| |||| ||||| |||||||||
Db         126 TATCCCAAAGATTCGGACAAACACATGCTCGCAAAACAAGCGGGACTAACCAGGAGCCAG 185

Qy       1462 GTGTCTAACTGGTTCATAAATGCTCGAGTTCGATTATGGAAGCCAATGGTAGAAGAAATG 1521
```

```
              ||||||||||||||||| || ||||| |||||  | ||||||||| ||||| |||||||||
Db        186 GTGTCTAACTGGTTCATCAACGCTCGGGTTCGCCTCTGGAAGCCCATGGTCGAAGAAATG 245

Qy       1522 TACTTGGAAGAAGTGAAGAATCAAGAACAAAACAGTACTAATA 1564
              ||||||||||||    ||||  | ||| |||     |  || | |
Db        246 TACTTGGAAGAGACCAAGAGCCGAGAGCAAGCTGGGTCTGAGA 288


RESULT 15
US-09-640-211A-1399/c
; Sequence 1399, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
;  APPLICANT: Wood, Marion
;  APPLICANT:  Shenk, Michael A.
;  APPLICANT:  McGrath, Annette
;  APPLICANT:  Glenn, Matthew
;  TITLE OF INVENTION: Compositions and Methods for the
;  TITLE OF INVENTION:  Modification of Gene Transcription
;  FILE REFERENCE: 11000.1021C1U
;  CURRENT APPLICATION NUMBER: US/09/640,211A
;  CURRENT FILING DATE:  2000-08-16
;  NUMBER OF SEQ ID NOS: 2368
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1399
;   LENGTH: 474
;   TYPE: DNA
;   ORGANISM: Eucalyptus grandis
US-09-640-211A-1399

  Query Match             3.9%;  Score 105.4;  DB 4;  Length 474;
  Best Local Similarity   77.9%;  Pred. No. 4e-17;
  Matches  127;  Conservative    0;  Mismatches   36;  Indels    0;  Gaps    0;

Qy       1402 TACCCAAAGGATTCAGACAAAATCATGCTTGCTAAGCAAACGGGGCTAACAAGGAGCCAG 1461
              || || || ||||| ||||||  |||||| || || ||| |||| ||||| |||||||||
Db        349 TATCCCAAAGATTCGGACAAACACATGCTCGCAAAACAAGCGGGACTAACCAGGAGCCAG 290

Qy       1462 GTGTCTAACTGGTTCATAAATGCTCGAGTTCGATTATGGAAGCCAATGGTAGAAGAAATG 1521
              ||||||||||||||||| || ||||| |||||  | ||||||||| ||||| |||||||||
Db        289 GTGTCTAACTGGTTCATCAACGCTCGGGTTCGCCTCTGGAAGCCCATGGTCGAAGAAATG 230

Qy       1522 TACTTGGAAGAAGTGAAGAATCAAGAACAAAACAGTACTAATA 1564
              ||||||||||||    ||||  | ||| |||     |  || | |
Db        229 TACTTGGAAGAGACCAAGAGCCGAGAGCAAGCTGGGTCTGAGA 187


Search completed: July 29, 2005, 14:20:30
Job time : 449 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 29, 2005, 08:10:18 ; Search time 1673 Seconds
(without alignments)
10580.930 Million cell updates/sec

Title: US-10-624-201A-1
Perfect score: 2735
Sequence: 1 catgcagagataaaaatata.....gcaaaaaaaaaaaaaaaaa 2735

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7287783 seqs, 3236178273 residues

Total number of hits satisfying chosen parameters: 14575566

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:*
22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:*
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	432.4	15.8	3033	18	US-10-424-599-130628	Sequence 130628,
2	392	14.3	1555	18	US-10-425-114-13353	Sequence 13353, A
3	378.2	13.8	1606	18	US-10-425-114-7804	Sequence 7804, Ap
4	353	12.9	1822	18	US-10-424-599-133558	Sequence 133558,
5	352.6	12.9	1892	18	US-10-424-599-63874	Sequence 63874, A
6	346.8	12.7	3311	18	US-10-424-599-141176	Sequence 141176,
7	324.6	11.9	2043	9	US-09-938-842A-1337	Sequence 1337, Ap
8	324.6	11.9	2043	11	US-09-938-842A-1337	Sequence 1337, Ap
9	324.6	11.9	2043	21	US-10-495-918-153	Sequence 153, App
10	324.6	11.9	2385	17	US-10-225-066A-335	Sequence 335, App
11	324.6	11.9	2385	17	US-10-374-780A-2671	Sequence 2671, Ap
12	324.6	11.9	2385	22	US-10-225-066A-335	Sequence 335, App
13	314	11.5	1716	18	US-10-425-114-9882	Sequence 9882, Ap
14	304	11.1	582	18	US-10-424-599-7028	Sequence 7028, Ap
15	295.6	10.8	3028	19	US-10-437-963-41007	Sequence 41007, A
16	278.8	10.2	2799	19	US-10-437-963-31585	Sequence 31585, A
17	268.8	9.8	1649	18	US-10-425-114-20970	Sequence 20970, A
18	266	9.7	677	19	US-10-767-701-3524	Sequence 3524, Ap
19	265.6	9.7	2943	20	US-10-425-115-118586	Sequence 118586,
20	264	9.7	831	19	US-10-767-701-10301	Sequence 10301, A
21	257.4	9.4	3029	19	US-10-437-963-60154	Sequence 60154, A
22	245.6	9.0	1645	18	US-10-424-599-65908	Sequence 65908, A
23	244.2	8.9	1637	18	US-10-425-114-13359	Sequence 13359, A
24	243.2	8.9	1546	18	US-10-425-114-29785	Sequence 29785, A
25	243.2	8.9	2736	18	US-10-424-599-37574	Sequence 37574, A
26	240	8.8	1498	18	US-10-425-114-33974	Sequence 33974, A
27	240	8.8	2313	18	US-10-425-114-15195	Sequence 15195, A
28	240	8.8	2970	20	US-10-425-115-97705	Sequence 97705, A
29	230.2	8.4	1452	18	US-10-425-114-5047	Sequence 5047, Ap
30	227	8.3	1452	20	US-10-425-115-118580	Sequence 118580,
31	226.2	8.3	1280	18	US-10-425-114-13349	Sequence 13349, A
32	225.4	8.2	1599	9	US-09-938-842A-2154	Sequence 2154, Ap
33	225.4	8.2	1599	11	US-09-938-842A-2154	Sequence 2154, Ap
34	225.4	8.2	2486	18	US-10-424-599-55280	Sequence 55280, A
35	224	8.2	1884	9	US-09-938-842A-905	Sequence 905, App
36	224	8.2	1884	11	US-09-938-842A-905	Sequence 905, App
37	224	8.2	1983	14	US-10-286-264-17	Sequence 17, Appl
38	224	8.2	1983	17	US-10-225-066A-819	Sequence 819, App
39	224	8.2	1983	17	US-10-374-780A-2481	Sequence 2481, Ap
40	224	8.2	1983	18	US-10-412-699B-637	Sequence 637, App
41	224	8.2	1983	22	US-10-225-066A-819	Sequence 819, App
42	222.6	8.1	2545	17	US-10-225-068-235	Sequence 235, App
43	222.6	8.1	2545	17	US-10-374-780A-2329	Sequence 2329, Ap
44	222.6	8.1	2545	18	US-10-412-699B-267	Sequence 267, App
45	222.6	8.1	2545	21	US-10-225-068-235	Sequence 235, App

ALIGNMENTS

RESULT 1
US-10-424-599-130628
; Sequence 130628, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 130628
; LENGTH: 3033
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_88966C.1
US-10-424-599-130628

Query Match 15.8%; Score 432.4; DB 18; Length 3033;
Best Local Similarity 72.4%; Pred. No. 3.8e-90;
Matches 575; Conservative 0; Mismatches 216; Indels 3; Gaps 1;

Qy 758 TTTTAGGCTCTAAGTATCTGAAAGCTGCACAAGAGCTTCTTGATGAAGTTGTTAATATTG 817
Db 861 TTATGGGGTCAAAGTACTTGAAAGCTGCACAGGAGCTTCTGGATGAAGTTGTGAACGTGG 920

```
Qy        818 TTGGAAAAAGCATCAAAGGAGATGATCAAAAGAAGGATAATTCAATGAATAAAGAATCAA 877
Db        921 GAAAAGGAATCTACAAAGAGGAGAAATTCTCAGAGAGCGTGAAAGCGAATAGGGAATCAA 980

Qy        878 TGCCTTTGGCTAGTGATGTCAACACTAATAGTTCTGGTGGTGGTGAAAGTAGCAGCAGGC 937
Db        981 CTAATTCTGGTGCTGCTGGTGATGGTGGAGATGGTTCAAGTGGTGGAGGAGAAAACAGTG 1040

Qy        938 AGAAAAATGAAGTTGCTGTTGAGCTTACAACTGCTCAAAGACAAGAACTTCAAATGAAAA 997
Db       1041 CAGGGAAACAAGT---GGTTGAACTCAGCACCGCTCAGAGGCAAGAGCTTCAGATGAAGA 1097

Qy        998 AAGCCAAGCTTCTTGCCATGCTTGAAGAGGTGGAGCAAAGGTACAGACAGTACCATCACC 1057
Db       1098 AGTCCAAGCTTGTGACCATGCTCGATGAGGTAGAACAAAGGTACAGACAATATCACCACC 1157

Qy       1058 AAATGCAAATAATTGTATTATCATTTGAGCAAGTAGCAGGAATTGGATCAGCCAAATCAT 1117
Db       1158 AAATGCAAATTGTGGTATCGTCATTTGAGCAAGCAGCAGGTTATGGGGCGGCAAAATCTT 1217

Qy       1118 ACACTCAATTAGCTTTGCATGCAATTTCGAAGCAATTCAGATGCCTAAAGGATGCAATTG 1177
Db       1218 ACACTGCCCTTGCCCTAAAAACTATCTCAAAGCAATTCAGGTGCCTGAAAGATGCAATCT 1277

Qy       1178 CTGAGCAAGTAAAGGCGACGAGCAAGAGTTTAGGTGAAGAGGAAGGCTTGGGAGGGAAAA 1237
Db       1278 CTGCACAAATCAAAGCCACGAGCAAGACATTGGGTGAAGATGATTGCTTAGGAGTTAAAG 1337

Qy       1238 TCGAAGGCTCAAGACTCAAATTTGTGGACCATCATCTAAGGCAACAACGCGCGCTGCAAC 1297
Db       1338 TAGAAGGGTCAAGGCTTAGGTATGTTGACCACCATCTCAGGCAACAAAGGGCACTTCAGC 1397

Qy       1298 AGATAGGAATGATGCAACCAAATGCTTGGAGACCCCAAAGAGGTTTACCTGAAAGAGCTG 1357
Db       1398 AGCTAGGAATGATTCAACCCAATGCATGGAGGCCCCAAAGAGGCTTGCCTGAACGAGCTG 1457

Qy       1358 TCTCTGTCCTTCGTGCTTGGCTTTTCGAGCATTTTCTTCATCCTTACCCAAAGGATTCAG 1417
Db       1458 TTTCCATTCTTCGGGCTTGGCTTTTTGAGCATTTCCTTCACCCTTATCCAAAGGACTCCG 1517

Qy       1418 ACAAAATCATGCTTGCTAAGCAAACGGGGCTAACAAGGAGCCAGGTGTCTAACTGGTTCA 1477
Db       1518 ATAAAGTTATGCTTGCTAAACAAACTGGACTTGCTAGGAGCCAGGTGTCAAACTGGTTTA 1577

Qy       1478 TAAATGCTCGAGTTCGATTATGGAAGCCAATGGTAGAAGAAATGTACTTGGAAGAAGTGA 1537
Db       1578 TCAATGCCCGAGTTCGGCTTTGGAAGCCAATGGTTGAAGAAATGTACTTGGAAGAAATCA 1637

Qy       1538 AGAATCAAGAACAA 1551
Db       1638 AGGAGCATGAACAA 1651
```

RESULT 2
US-10-425-114-13353
; Sequence 13353, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 13353
; LENGTH: 1555
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: JC-GMST02400042B10_FLI
US-10-425-114-13353

Query Match 14.3%; Score 392; DB 18; Length 1555;
Best Local Similarity 78.4%; Pred. No. 7.4e-81;
Matches 468; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

```
Qy        955 GTTGAGCTTACAACTGCTCAAAGACAAGAACTTCAAATGAAAAAAGCCAAGCTTCTTGCC 1014
Db          92 GTTGAACTCAGCACCGCTCAGAGGCAAGAGCTTCAGATGAAGAAGTCCAAGCTTGTGACC 151

Qy       1015 ATGCTTGAAGAGGTGGAGCAAAGGTACAGACAGTACCATCACCAAATGCAAATAATTGTA 1074
Db        152 ATGCTCGATGAGGTAGAACAAAGGTACAGACAATATCACCACCAAATGCAAATTGTGGTA 211

Qy       1075 TTATCATTTGAGCAAGTAGCAGGAATTGGATCAGCCAAATCATACACTCAATTAGCTTTG 1134
Db        212 TCGTCATTTGAGCAGGCAGCAGGTTATGGGGCGGCAAAATCTTACACTGCCCTTGCCCTA 271

Qy       1135 CATGCAATTTCGAAGCAATTCAGATGCCTAAAGGATGCAATTGCTGAGCAAGTAAAGGCG 1194
Db        272 AAGACTATCTCAAAGCAATTCAGGTGCCTTAAAGATGCAATCTCTGCACAAATCAAAGCC 331

Qy       1195 ACGAGCAAGAGTTTAGGTGAAGAGGAAGGCTTGGGAGGGAAAATCGAAGGCTCAAGACTC 1254
Db        332 ACGAGCAAGACATTGGGTGAAGATGATTGTTTAGGAGTTAAAGTAGAAGGGTCGAGGCTT 391

Qy       1255 AAATTTGTGGACCATCATCTAAGGCAACAACGCGCGCTGCAACAGATAGGAATGATGCAA 1314
Db        392 AGGTTTGTTGACCACCATCTCAGGCAACAAAGGGCACTTCAGCAGCTAGGAATGATTCAA 451

Qy       1315 CCAAATGCTTGGAGACCCCAAAGAGGTTTACCTGAAAGAGCTGTCTCTGTCCTTCGTGCT 1374
Db        452 CCCAATGCATGGAGGCCCCAAAGAGGCTTGCCTGAACGAGCTGTTTCCATTCTTCGGGCT 511

Qy       1375 TGGCTTTTCGAGCATTTTCTTCATCCTTACCCAAAGGATTCAGACAAAATCATGCTTGCT 1434
Db        512 TGGCTTTTTGAGCATTTCCTTCACCCTTATCCAAAGGACTCCGATAAAGTTATGCTTGCT 571

Qy       1435 AAGCAAACGGGGCTAACAAGGAGCCAGGTGTCTAACTGGTTCATAAATGCTCGAGTTCGA 1494
Db        572 AAACAAACTGGACTTGCTAGGAGCCAGGTGTCAAACTGGTTTATCAATGCCCGAGTTCGG 631

Qy       1495 TTATGGAAGCCAATGGTAGAAGAAATGTACTTGGAAGAAGTGAAGAATCAAGAACAA 1551
Db        632 CTTTGGAAGCCAATGGTTGAAGAAATGTACTTGGAAGAAATCAAGGAGCATGAACAA 688
```

RESULT 3
US-10-425-114-7804
; Sequence 7804, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 7804
; LENGTH: 1606
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:

```
;   OTHER INFORMATION: Clone ID: 700684286_FLI
US-10-425-114-7804

  Query Match             13.8%;  Score 378.2;  DB 18;  Length 1606;
  Best Local Similarity   73.6%;  Pred. No. 1.3e-77;
  Matches  482;  Conservative    0;  Mismatches  173;  Indels    0;  Gaps    0;

Qy        908 GTTCTGGTGGTGGTGAAAGTAGCAGCAGGCAGAAAAATGAAGTTGCTGTTGAGCTTACAA 967
              ||   || || |   | | ||  |  |   |   |   ||   | ||||| || | |
Db         10 GTGGAGGAGGGGAGAATAATAATAATAATGATGGAGGCAAACAAGGAGTTGAACTCAGCA 69

Qy        968 CTGCTCAAAGACAAGAACTTCAAATGAAAAAAGCCAAGCTTCTTGCCATGCTTGAAGAGG 1027
              | || || |||||||| ||||| ||||| ||  | || ||| |  |||||| || ||||
Db         70 CAGCACAGAGACAAGAGCTTCAGATGAAGAAGTCAAAACTTGTGACCATGCTAGATGAGG 129

Qy       1028 TGGAGCAAAGGTACAGACAGTACCATCACCAAATGCAAATAATTGTATTATCATTTGAGC 1087
              |||||||||||||| ||||||| || ||||||||||||  |  ||  ||||||||||
Db        130 TGGAGCAAAGGTACCGACAGTATCACCACCAAATGCAAGTTGTGATAACATCATTTGAGC 189

Qy       1088 AAGTAGCAGGAATTGGATCAGCCAAATCATACACTCAATTAGCTTTGCATGCAATTTCGA 1147
              ||| ||| ||  ||||| | || || || |||||      | || ||  |  |||| || |
Db        190 AAGCAGCGGGTGTTGGAGCTGCAAAGTCTTACACAGCCCTTGCGTTAAAGACAATCTCAA 249

Qy       1148 AGCAATTCAGATGCCTAAAGGATGCAATTGCTGAGCAAGTAAAGGCGACGAGCAAGAGTT 1207
              |||||||  | || || || |||||||   ||    ||| ||||| | |||||||  |  |
Db        250 AGCAATTTCGGTGTCTCAAAGATGCAATCTCTTCACAAATAAAGACCACGAGCAAAACCT 309

Qy       1208 TAGGTGAAGAGGAAGGCTTGGGAGGGAAAATCGAAGGCTCAAGACTCAAATTTGTGGACC 1267
              |||||||||| ||  |||||||||  ||  | ||||| || || || |  | ||| || |
Db        310 TAGGTGAAGATGATTGCTTGGGAGTTAAGGTAGAAGGTTCGAGGCTTAGGTATGTTGATC 369

Qy       1268 ATCATCTAAGGCAACAACGCGCGCTGCAACAGATAGGAATGATGCAACCAAATGCTTGGA 1327
              |||| || || |||||||| ||  | ||||| | |||||||| ||||  ||||||||||
Db        370 ATCAACTGAGACAACAACGTGCATTACAACAGCTTGGAATGATCCAACACAATGCTTGGA 429

Qy       1328 GACCCCAAAGAGGTTTACCTGAAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTTTTCGAGC 1387
              | |||||||||| || |||||| | ||||| |||| |||||||||||  || |
Db        430 GGCCCCAAAGAGGCTTGCCTGAACGTGCTGTTTCTGTTCTTCGAGCTTGGCTTTTTGAAC 489

Qy       1388 ATTTTCTTCATCCTTACCCAAAGGATTCAGACAAAATCATGCTTGCTAAGCAAACGGGGC 1447
              ||||  | || ||||| || ||||| || || ||  | |||||||| || ||||| ||||
Db        490 ATTTCTTGCACCCTTATCCTAAGGACTCCGATAAGGTTATGCTTGCAAAACAAACTGGAC 549

Qy       1448 TAACAAGGAGCCAGGTGTCTAACTGGTTCATAAATGCTCGAGTTCGATTATGGAAGCCAA 1507
              | ||  |||||||||||||||||||||| |||||||| ||||||||  ||||||||||||
Db        550 TTACTCGGAGCCAGGTGTCTAACTGGTTTATAAATGCCCGAGTTCGGCTATGGAAGCCAA 609

Qy       1508 TGGTAGAAGAAATGTACTTGGAAGAAGTGAAGAATCAAGAACAAAACAGTACTAA 1562
              |||||||||||||||||||||||||| || ||  |  ||  | | ||| ||  || |
Db        610 TGGTAGAAGAAATGTACTTGGAAGAGGTAAAACAAGAACCAAACAATAGCTCTCA 664


RESULT 4
US-10-424-599-133558
; Sequence 133558, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 133558
;   LENGTH: 1822
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_91608C.1
US-10-424-599-133558

  Query Match             12.9%;  Score 353;  DB 18;  Length 1822;
  Best Local Similarity   72.6%;  Pred. No. 1.1e-71;
  Matches  471;  Conservative    0;  Mismatches  175;  Indels    3;  Gaps    1;

Qy        908 GTTCTGGTGGTGGTGAAAGTAGCAGCAGGCAGAAAAATGAAGTTGCTGTTGAGCTTACAA 967
              ||   || || |   | | ||  |  |   |   |   ||   | ||||| || | |
Db         19 GTGGAGGAGGGGAGAATAATAATAATAATGATGGAGGCAAACAAGGAGTTGAACTCAGCA 78

Qy        968 CTGCTCAAAGACAAGAACTTCAAATGAAAAAAGCCAAGCTTCTTGCCATGCTTGAAGAGG 1027
              | || || |||||||| ||||| ||||| ||  | || ||| |  |||||| || ||||
Db         79 CAGCACAGAGACAAGAGCTTCAGATGAAGAAGTCAAAACTTGTGACCATGCTAGATGAGG 138

Qy       1028 TGGAGCAAAGGTACAGACAGTACCATCACCAAATGCAAATAATTGTATTATCATTTGAGC 1087
              |||||||||| || |||||||| || ||||||||||| |  |  ||  ||||||||||
Db        139 TGGAGCAAAGGGACCGACAGTATCACCACCAAATGCAAGTTGTGATAACATCATTTGAGC 198

Qy       1088 AAGTAGCAGGAATTGGATCAGCCAAATCATACACTCAATTAGCTTTGCATGCAATTTCGA 1147
              ||| ||| ||  ||||| | || || || |||||      | || ||  |  |||| || |
Db        199 AAGCAGCGGGTGTTGGAGCTGCAAAGTCTTACACAGCCCTTGCGTTAAAGACAATCTCAA 258

Qy       1148 AGCAATTCAGATGCCTAAAGGATGCAATTGCTGAGCAAGTAAAGGCGACGAGCAAGAGTT 1207
              |||||||  | || || || |||||||| ||    ||| | ||| | || |||||||  |
Db        259 AGCAATTTCGGTGTCTCAAAGATGCAATCTCTTCACAAATCAAGACAACTAGCAAGACCT 318

Qy       1208 TAGGTGAAGAGGAAGGCTTGGGAGGGAAAATCGAAGGCTCAAGACTCAAATTTGTGGACC 1267
              ||||||||||  |  || ||||||  ||  | || || || || || |  | ||| || |
Db        319 TAGGTGAAGATAATTGCCTGGGAGTTAAGGTAGAGGGTTCGAGGCTAAGGTATGTTGATC 378

Qy       1268 ATCATCTAAGGCAACAACGCGCGCTGCAACAGATAGGAATGATGCAACCAAATGCTTGGA 1327
              |||| |  || |||||||| ||  |  |||| | |||||||| ||||  || |||||||
Db        379 ATCAACAGAGACAACAACGTGCATT---ACAGCTTGGAATGATCCAACACAACGCTTGGA 435

Qy       1328 GACCCCAAAGAGGTTTACCTGAAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTTTTCGAGC 1387
              | |||||||||| || |||||| | ||||||||||||||||||||||||||||| || |
Db        436 GGCCCCAAAGAGGCTTGCCTGAACGTGCTGTCTCTGTCCTTCGTGCTTGGCTTTTTGAAC 495

Qy       1388 ATTTTCTTCATCCTTACCCAAAGGATTCAGACAAAATCATGCTTGCTAAGCAAACGGGGC 1447
              ||||  | || || || || ||||| || || ||  | |||||||| || ||||| ||||
Db        496 ATTTCTTGCACCCCTATCCTAAGGACTCCGATAAGGTTATGCTTGCAAAACAAACTGGGC 555

Qy       1448 TAACAAGGAGCCAGGTGTCTAACTGGTTCATAAATGCTCGAGTTCGATTATGGAAGCCAA 1507
              | ||  |||||||||| |||||||||| |||||||| ||||||||  ||||||||||||
Db        556 TTACTCGGAGCCAGGTATCTAACTGGTTTATAAATGCCCGAGTTCGGCTATGGAAGCCAA 615

Qy       1508 TGGTAGAAGAAATGTACTTGGAAGAAGTGAAGAATCAAGAACAAAACAG 1556
              |||||||||||||||||||||||||| || ||  |  |  | | || ||
Db        616 TGGTAGAAGAAATGTACTTGGAAGAGGTAAAACAAGAGCCAAACAATAG 664


RESULT 5
US-10-424-599-63874
; Sequence 63874, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 63874
;   LENGTH: 1892
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_28692C.1
US-10-424-599-63874
```

Query Match 12.9%; Score 352.6; DB 18; Length 1892;
Best Local Similarity 71.1%; Pred. No. 1.4e-71;
Matches 466; Conservative 0; Mismatches 189; Indels 0; Gaps 0;

```
Qy        908 GTTCTGGTGGTGGTGAAAGTAGCAGCAGGCAGAAAAATGAAGTTGCTGTTGAGCTTACAA 967
              ||   || || |   | | ||  |  |   |   |    ||   |  ||||| || |  |
Db         32 GTGGAGGAGGGGAGAATAATAATAATAATGATGGAGGCAAACAAGGAGTTGAACTCAGCA 91

Qy        968 CTGCTCAAAGACAAGAACTTCAAATGAAAAAAGCCAAGCTTCTTGCCATGCTTGAAGAGG 1027
              | || || |||||||| ||||| ||||| ||  | || |||  |  |||||| || ||||
Db         92 CAGCACAGAGACAAGAGCTTCAGATGAAGAAGTCAAAACTTGTGACCATGCTAGATGAGG 151

Qy        1028 TGGAGCAAAGGTACAGACAGTACCATCACCAAATGCAAATAATTGTATTATCATTTGAGC 1087
               |||||||||||||| ||||||| || |||||||||||| |  |  ||  ||||||||||
Db         152 TGGAGCAAAGGTACCGACAGTATCACCACCAAATGCAAGTTGTGATAACATCATTTGAGC 211

Qy        1088 AAGTAGCAGGAATTGGATCAGCCAAATCATACACTCAATTAGCTTTGCATGCAATTTCGA 1147
               ||| ||| ||  ||||| | || || || |||||     | || ||  |  |||| || |
Db         212 AAGCAGCGGGTGTTGGAGCTGCAAAGTCTTACACAGCCCTTGCGTTAAAGACAATCTCAA 271

Qy        1148 AGCAATTCAGATGCCTAAAGGATGCAATTGCTGAGCAAGTAAAGGCGACGAGCAAGAGTT 1207
               |||||||  | || || || |||||||  ||   ||| |||||  | |||||||| |  |
Db         272 AGCAATTTCGGTGTCTCAAAGATGCAATCTCTTCACAAATAAAGACCACGAGCAAAACCT 331

Qy        1208 TAGGTGAAGAGGAAGGCTTGGGAGGGAAAATCGAAGGCTCAAGACTCAAATTTGTGGACC 1267
               |||||||||| ||  |||||||||  ||  | ||||| || || || |  | ||| || |
Db         332 TAGGTGAAGATGATTGCTTGGGAGTTAAGGTAGAAGGTTCGAGGCTTAGGTATGTTGATC 391

Qy        1268 ATCATCTAAGGCAACAACGCGCGCTGCAACAGATAGGAATGATGCAACCAAATGCTTGGA 1327
               |||| || || |||||||| ||  | |||||| | ||||||  |   |          |
Db         392 ATCAACTGAGACAACAACGTGCATTACAACAGCTTGGAATGCGGACGCGTGGGCGGACGC 451

Qy        1328 GACCCCAAAGAGGTTTACCTGAAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTTTTCGAGC 1387
               |   |||||||| || |||||| | ||||| ||||| ||||| ||||||||||| || |
Db         452 GTGGGCAAAGAGGCTTGCCTGAACGTGCTGTTTCTGTTCTTCGAGCTTGGCTTTTTGAAC 511

Qy        1388 ATTTTCTTCATCCTTACCCAAAGGATTCAGACAAAATCATGCTTGCTAAGCAAACGGGGC 1447
               ||||  | || ||||| || ||||| || || ||  | |||||||| || ||||| || |
Db         512 ATTTCTTGCACCCTTATCCTAAGGACTCCGATAAGGTTATGCTTGCAAAACAAACTGGAC 571

Qy        1448 TAACAAGGAGCCAGGTGTCTAACTGGTTCATAAATGCTCGAGTTCGATTATGGAAGCCAA 1507
               | ||  |||||||||||||||||||||| |||||||| ||||||||  ||||||||||||
Db         572 TTACTCGGAGCCAGGTGTCTAACTGGTTTATAAATGCCCGAGTTCGGCTATGGAAGCCAA 631

Qy        1508 TGGTAGAAGAAATGTACTTGGAAGAAGTGAAGAATCAAGAACAAAACAGTACTAA 1562
               |||||||||||||||||||||||||| || ||  |  ||  | | || ||  || |
Db         632 TGGTAGAAGAAATGTACTTGGAAGAGGTAAAACAAGAACCAAACAATAGCTCTCA 686
```

RESULT 6
US-10-424-599-141176
; Sequence 141176, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 141176
; LENGTH: 3311
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_98494C.1
US-10-424-599-141176

Query Match 12.7%; Score 346.8; DB 18; Length 3311;
Best Local Similarity 70.6%; Pred. No. 4.3e-70;
Matches 476; Conservative 0; Mismatches 197; Indels 1; Gaps 1;

```
Qy         957 TGAGCTTACAACTGCTCAAAGACAAGAACTTCAAATGAAAAAAGCCAAGCTTCTTGCCAT 1016
               |||  |  ||||| |  |||||||||| |||| || || ||||| |||||  |   |||
Db        1473 TGATTTGTCAACTACGGAAAGACAAGAAATTCAGATAAAGAAAGCAAAGCTAATCAACAT 1532

Qy        1017 GCTTGAAGAGGTGGAGCAAAGGTACAGACAGTA-CCATCACCAAATGCAAATAATTGTAT 1075
               |||||| |||||||||||||||||||||| |||||  ||  |||| ||| | ||| |  |||
Db        1533 GCTTGATGAGGTGGAGCAAAGGTACAGGCAGTATTCACAACCAGATGAAGATAGTGATAT 1592

Qy        1076 TATCATTTGAGCAAGTAGCAGGAATTGGATCAGCCAAATCATACACTCAATTAGCTTTGC 1135
                ||||||||||||||  || ||||| || ||||| | | | ||||||     | || ||||
Db        1593 CCTCATTTGAGCAAGCTGCTGGAATAGGCTCAGCAAGAACGTACACTGCGCTTGCATTGC 1652

Qy        1136 ATGCAATTTCGAAGCAATTCAGATGCCTAAAGGATGCAATTGCTGAGCAAGTAAAGGCGA 1195
               |  |||||||||||||||  | ||  | ||||||||||  | | |||| | |  ||
Db        1653 AAACAATTTCGAAGCAATTTCGGTGTTTGAAGGATGCAATAACGGGGCAAATCAGAGCTG 1712

Qy        1196 CGAGCAAGAGTTTAGGTGAAGAGGAAGGCTTGGGAGGGAAAATCGAAGGCTCAAGACTCA 1255
               | | ||||||||| || |||||||  | || || |  ||||| |||||  || || ||||
Db        1713 CAAACAAGAGTTTGGGAGAAGAGGATTGTTTTGGTGCTAAAATTGAAGGATCTAGGCTCA 1772

Qy        1256 AATTTGTGGACCATCATCTAAGGCAACAACGCGCGCTGCAACAGATAGGAATGATGCAAC 1315
               ||| |||||||||||||||||| ||||  | ||  | ||| ||   |  |||||||  || |
Db        1773 AATATGTGGACCATCATCTAAGGTAACAGAGGGCTTTTCAGCACCTGAGAATGATCCATC 1832

Qy        1316 CAAATGCTTGGAGACCCCAAAGAGGTTTACCTGAAAGAGCTGTCTCTGTCCTTCGTGCTT 1375
                |||||||||||||||   |||   |  |   | |||  | |||||  || | ||  ||
Db        1833 ACAATGCTTGGAGACCCCTGCGAGTACTGGCCAGGATAGCACTTTCTGTTCTGCCTGATT 1892

Qy        1376 GGCTTTTCGAGCATTTTCTTCATCCTTACCCAAAGGATTCAGACAAAATCATGCTTGCTA 1435
               ||   || || || |||| | |||||||| | ||||| || |||||   ||||||||||
Db        1893 GGAAGTTTGAACACTTTCATAATCCTTACACCAAGGACTCGGACAAGCACATGCTTGCTA 1952

Qy        1436 AGCAAACGGGGCTAACAAGGAGCCAGGTGTCTAACTGGTTCATAAATGCTCGAGTTCGAT 1495
               | ||||| ||||| || |||||||||| || || |||||||| ||||||  |||||||
Db        1953 AACAAACAGGGCTTACCAGGAGCCAGGTTTCAAATTGGTTCATTAATGCTAGAGTTCGGC 2012

Qy        1496 TATGGAAGCCAATGGTAGAAGAAATGTACTTGGAAGAAGTGAAGAATCAAGAACAAAACA 1555
               | |||||||||||||| || |||||||||||||  ||  ||||  | || |||| |||
Db        2013 TTTGGAAGCCAATGGTGGAGGAAATGTACTTGGAGGAGATGAAAGAGCATGAACTAAATG 2072

Qy        1556 GTACTAATACTTCAGGAGATAACAAAAACAAAGAGACCAATATAAGTGCTCCAAATGAAG 1615
               || || |      |  |   || |      ||||  |   || ||   |  ||| |  |
Db        2073 GTTCTGAGGAGAAATCAAGCAAGAGTGGTGAAGATCCTGCTACAAAAACCACAAGTCCAC 2132

Qy        1616 AGAAACATCCAATT 1629
               |  |  | |  | |
Db        2133 AAGAGAAACGTACT 2146
```

RESULT 7
US-09-938-842A-1337
; Sequence 1337, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIP1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22

; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1337
; LENGTH: 2043
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1337

Query Match 11.9%; Score 324.6; DB 9; Length 2043;
Best Local Similarity 65.2%; Pred. No. 5.1e-65;
Matches 534; Conservative 0; Mismatches 264; Indels 21; Gaps 3;

```
Qy        758 TTTTAGGCTCTAAGTATCTGAAAGCTGCACAAGAGCTTCTTGATGAAGTTGTTAATATTG 817
              || |  |||| |||||  |||| || ||||||||||||||||| ||||| || ||   ||
Db        569 TTGTTAGCTCCAAGTACTTGAAGGCAGCACAAGAGCTTCTTGACGAAGTAGTCAACGCTG 628

Qy        818 TTGGAAAAAGCATCAAAGGAGATGATCAAAAGAAGGATAATTCAATGAATAAAGAATCAA 877
              |   |   ||| || |  |  |||             || |  ||     ||
Db        629 ATTCCGATGACATGAACGCTAAATCCCAACTATTCTCATCGAAAAAGGGTAGTTGCGGAA 688

Qy        878 TGCCTTTGGCTAGTGATGTCAACACTAATAGTTCTGGTGGTGGTGAAAGTAGCAGCAGGC 937
                  ||    |  |       |     | |||| || |  |       |  || |
Db        689 ATGATAAACCTGTCGGAGAATCATCGGCCGGCGCTGGAGGAGAAGGTTCCGGTGGCGGAG 748

Qy        938 AGAAAAATG------AAGTTGCTGTTGAGCTTACAACTGCTCAAAGACAAGAACTTCAAA 991
                ||     |      ||  | | ||  |||||    || ||  | |||||||| | || |
Db        749 CAGAAGCAGCCGGGAAACGTCCGGTGGAGCTAGGCACGGCAGAGAGACAAGAAATACAGA 808

Qy        992 TGAAAAAAGCCAAGCTTCTTGCCATGCTTGAAGAGGTGGAGCAAAGGTACAGACAGTACC 1051
              |||| ||||| || |||  |  ||||||| | |||||||||||| || || ||||||||||
Db        809 TGAAGAAAGCAAAACTTAGTAACATGCTTCATGAGGTGGAGCAGAGATATAGACAGTACC 868

Qy       1052 ATCACCAAATGCAAATAATTGTATTATCATTTGAGCAAGTAGCAGGAATTGGATCAGCCA 1111
              | || || ||||| ||  |  | |  || || |||||||  |||||| || |||||||| |
Db        869 ACCAGCAGATGCAGATGGTGATCTCTTCGTTCGAGCAAGCGGCAGGGATAGGATCAGCGA 928

Qy       1112 AATCATACACTCAATTAGCTTTGCATGCAATTTCGAAGCAATTCAGATGCCTAAAGGATG 1171
              | ||||||||     |||| ||| |  | || || |  || ||| | ||| | || || | |
Db        929 AGTCATACACGTCGCTAGCATTGAAAACCATATCAAGACAGTTCCGTTGCTTGAAAGAGG 988

Qy       1172 CAATTGCTGAGCAAGTAAAGGCGACGAGCAAGAGTTTAGGTGAAGAGGAAGGCTTG---- 1227
              | || ||||  ||  |||| ||| | | ||||||| | || || || ||    ||
Db        989 CGATCGCTGGTCAGATAAAAGCGGCCAACAAGAGTCTTGGGGAGGAAGATTCAGTGTCTG 1048

Qy       1228 --GGAGGGAAAATCGAAGGCTCAAGACTCAAATTTGTGGACCATCATCTAAGGCAACAAC 1285
                |  ||||   | || || || || ||||| || |||||||| ||  | || || |||
Db       1049 GTGTTGGGAGGTTTGAGGGGTCGAGGCTCAAGTTCGTGGACCACCACTTGAGACAGCAAA 1108

Qy       1286 GCGCGCTGCAACAGATAGGAATGATGCAAC---------CAAATGCTTGGAGACCCCAAA 1336
              | || || |||||  | |||||||| ||||          ||||||||||||| |||
Db       1109 GAGCTCTTCAACAACTGGGAATGATTCAACATCCTTCCAATAATGCTTGGAGACCTCAAC 1168

Qy       1337 GAGGTTTACCTGAAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTTTTCGAGCATTTTCTTC 1396
              | ||| | || ||| |||| ||||| || || |||||||||| ||||| || |||||||
Db       1169 GTGGTCTCCCAGAACGAGCCGTCTCAGTTCTCCGTGCTTGGCTCTTCGAACACTTTCTTC 1228

Qy       1397 ATCCTTACCCAAAGGATTCAGACAAAATCATGCTTGCTAAGCAAACGGGGCTAACAAGGA 1456
              |||| ||||| |||||||| |||||   |||||| |||||||||| || || ||  | |
Db       1229 ATCCATACCCTAAGGATTCGGACAAGCACATGCTAGCTAAGCAAACAGGACTCACTCGTA 1288

Qy       1457 GCCAGGTGTCTAACTGGTTCATAAATGCTCGAGTTCGATTATGGAAGCCAATGGTAGAAG 1516
              |||||||||| |||||||| |||| ||   |||||| |||||||| |||||||| || |
Db       1289 GCCAGGTGTCGAACTGGTTTATAAACGCGAGAGTTCGGTTATGGAAACCAATGGTGGAGG 1348

Qy       1517 AAATGTACTTGGAAGAAGTGAAGAATCAAGAACAAAACA 1555
              | |||||| |||| ||| ||||| | || | | | ||||
Db       1349 AGATGTACATGGAGGAAATGAAGGAGCAGGCAAAGAACA 1387
```

RESULT 8
US-09-938-842A-1337
; Sequence 1337, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIP1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1337
; LENGTH: 2043
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1337

Query Match 11.9%; Score 324.6; DB 11; Length 2043;
Best Local Similarity 65.2%; Pred. No. 5.1e-65;
Matches 534; Conservative 0; Mismatches 264; Indels 21; Gaps 3;

```
Qy        758 TTTTAGGCTCTAAGTATCTGAAAGCTGCACAAGAGCTTCTTGATGAAGTTGTTAATATTG 817
              || |  |||| |||||  |||| || ||||||||||||||||| ||||| || ||   ||
Db        569 TTGTTAGCTCCAAGTACTTGAAGGCAGCACAAGAGCTTCTTGACGAAGTAGTCAACGCTG 628

Qy        818 TTGGAAAAAGCATCAAAGGAGATGATCAAAAGAAGGATAATTCAATGAATAAAGAATCAA 877
              |   |   ||| || |  |  |||             || |  ||     ||
Db        629 ATTCCGATGACATGAACGCTAAATCCCAACTATTCTCATCGAAAAAGGGTAGTTGCGGAA 688

Qy        878 TGCCTTTGGCTAGTGATGTCAACACTAATAGTTCTGGTGGTGGTGAAAGTAGCAGCAGGC 937
                  ||    |  |       |     | |||| || |  |       |  || |
Db        689 ATGATAAACCTGTCGGAGAATCATCGGCCGGCGCTGGAGGAGAAGGTTCCGGTGGCGGAG 748

Qy        938 AGAAAAATG------AAGTTGCTGTTGAGCTTACAACTGCTCAAAGACAAGAACTTCAAA 991
                ||     |      ||  | | ||  |||||    || ||  | |||||||| | || |
Db        749 CAGAAGCAGCCGGGAAACGTCCGGTGGAGCTAGGCACGGCAGAGAGACAAGAAATACAGA 808

Qy        992 TGAAAAAAGCCAAGCTTCTTGCCATGCTTGAAGAGGTGGAGCAAAGGTACAGACAGTACC 1051
              |||| ||||| || |||  |  ||||||| | |||||||||||| || || ||||||||||
Db        809 TGAAGAAAGCAAAACTTAGTAACATGCTTCATGAGGTGGAGCAGAGATATAGACAGTACC 868

Qy       1052 ATCACCAAATGCAAATAATTGTATTATCATTTGAGCAAGTAGCAGGAATTGGATCAGCCA 1111
              | || || ||||| ||  |  | |  || || |||||||  |||||| || |||||||| |
Db        869 ACCAGCAGATGCAGATGGTGATCTCTTCGTTCGAGCAAGCGGCAGGGATAGGATCAGCGA 928

Qy       1112 AATCATACACTCAATTAGCTTTGCATGCAATTTCGAAGCAATTCAGATGCCTAAAGGATG 1171
              | ||||||||     |||| ||| |  | || || |  || ||| | ||| | || || | |
Db        929 AGTCATACACGTCGCTAGCATTGAAAACCATATCAAGACAGTTCCGTTGCTTGAAAGAGG 988

Qy       1172 CAATTGCTGAGCAAGTAAAGGCGACGAGCAAGAGTTTAGGTGAAGAGGAAGGCTTG---- 1227
              | || ||||  ||  |||| ||| | | ||||||| | || || || ||    ||
Db        989 CGATCGCTGGTCAGATAAAAGCGGCCAACAAGAGTCTTGGGGAGGAAGATTCAGTGTCTG 1048

Qy       1228 --GGAGGGAAAATCGAAGGCTCAAGACTCAAATTTGTGGACCATCATCTAAGGCAACAAC 1285
                |  ||||   | || || || || ||||| || |||||||| ||  | || || |||
Db       1049 GTGTTGGGAGGTTTGAGGGGTCGAGGCTCAAGTTCGTGGACCACCACTTGAGACAGCAAA 1108

Qy       1286 GCGCGCTGCAACAGATAGGAATGATGCAAC---------CAAATGCTTGGAGACCCCAAA 1336
              | || || |||||  | |||||||| ||||          ||||||||||||| |||
Db       1109 GAGCTCTTCAACAACTGGGAATGATTCAACATCCTTCCAATAATGCTTGGAGACCTCAAC 1168

Qy       1337 GAGGTTTACCTGAAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTTTTCGAGCATTTTCTTC 1396
              | ||| | || ||| |||| ||||| || || |||||||||| ||||| || |||||||
Db       1169 GTGGTCTCCCAGAACGAGCCGTCTCAGTTCTCCGTGCTTGGCTCTTCGAACACTTTCTTC 1228

Qy       1397 ATCCTTACCCAAAGGATTCAGACAAAATCATGCTTGCTAAGCAAACGGGGCTAACAAGGA 1456
              |||| ||||| |||||||| |||||   |||||| |||||||||| || || ||  | |
Db       1229 ATCCATACCCTAAGGATTCGGACAAGCACATGCTAGCTAAGCAAACAGGACTCACTCGTA 1288
```

```
Qy       1457 GCCAGGTGTCTAACTGGTTCATAAATGCTCGAGTTCGATTATGGAAGCCAATGGTAGAAG 1516
             |||||||||| |||||||| ||||| ||  ||||||| |||||||| |||||||| || |
Db       1289 GCCAGGTGTCGAACTGGTTTATAAACGCGAGAGTTCGGTTATGGAAACCAATGGTGGAGG 1348

Qy       1517 AAATGTACTTGGAAGAAGTGAAGAATCAAGAACAAAACA 1555
              | |||||| |||| ||| ||||| | || | | | ||||
Db       1349 AGATGTACATGGAGGAAATGAAGGAGCAGGCAAAGAACA 1387


RESULT 9
US-10-495-918-153
; Sequence 153, Application US/10495918
; Publication No. US20050009187A1
; GENERAL INFORMATION:
;  APPLICANT: RIKEN
;  TITLE OF INVENTION: A STRESS-RESPONSIVE PROMOTER AND A GENE CORDING FOR A STRESS-RESP
;  TITLE OF INVENTION:  TRANSCRIPTION FACTOR
;  FILE REFERENCE: RFH14-024T
;  CURRENT APPLICATION NUMBER: US/10/495,918
;  CURRENT FILING DATE:  2004-05-18
;  PRIOR APPLICATION NUMBER: JP 2001-353038
;  PRIOR FILING DATE: 2001-11-19
;  PRIOR APPLICATION NUMBER: JP 2002-20329
;  PRIOR FILING DATE: 2002-01-29
;  NUMBER OF SEQ ID NOS: 166
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 153
;   LENGTH: 2043
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(2040)
US-10-495-918-153

  Query Match            11.9%;  Score 324.6;  DB 21;  Length 2043;
  Best Local Similarity   65.2%;  Pred. No. 5.1e-65;
  Matches  534;  Conservative    0;  Mismatches  264;  Indels   21;  Gaps    3;

Qy        758 TTTTAGGCTCTAAGTATCTGAAAGCTGCACAAGAGCTTCTTGATGAAGTTGTTAATATTG 817
              || |  |||| |||||  |||| || ||||||||||||||||||| ||||| || ||   ||
Db        569 TTGTTAGCTCCAAGTACTTGAAGGCAGCACAAGAGCTTCTTGACGAAGTAGTCAACGCTG 628

Qy        818 TTGGAAAAAGCATCAAAGGAGATGATCAAAAGAAGGATAATTCAATGAATAAAGAATCAA 877
               |     |  ||| || |  |   |||                || |  ||     ||
Db        629 ATTCCGATGACATGAACGCTAAATCCCAACTATTCTCATCGAAAAAGGGTAGTTGCGGAA 688

Qy        878 TGCCTTTGGCTAGTGATGTCAACACTAATAGTTCTGGTGGTGGTGAAAGTAGCAGCAGGC 937
                 |   ||   |  |       |      |  |||| || |  |       |  || |
Db        689 ATGATAAACCTGTCGGAGAATCATCGGCCGGCGCTGGAGGAGAAGGTTCCGGTGGCGGAG 748

Qy        938 AGAAAAATG------AAGTTGCTGTTGAGCTTACAACTGCTCAAAGACAAGAACTTCAAA 991
                ||   |      ||  | | || |||||    || ||  | |||||||||| | || |
Db        749 CAGAAGCAGCCGGGAAACGTCCGGTGGAGCTAGGCACGGCAGAGAGACAAGAAATACAGA 808

Qy        992 TGAAAAAAGCCAAGCTTCTTGCCATGCTTGAAGAGGTGGAGCAAAGGTACAGACAGTACC 1051
              |||| ||||| || |||  |  ||||||| | |||||||||| || || ||||||||||
Db        809 TGAAGAAAGCAAAACTTAGTAACATGCTTCATGAGGTGGAGCAGAGATATAGACAGTACC 868

Qy       1052 ATCACCAAATGCAAATAATTGTATTATCATTTGAGCAAGTAGCAGGAATTGGATCAGCCA 1111
              | || || ||||| ||  |  | |  || || |||||||  ||||| || ||||||||| |
Db        869 ACCAGCAGATGCAGATGGTGATCTCTTCGTTCGAGCAAGCGGCAGGGATAGGATCAGCGA 928

Qy       1112 AATCATACACTCAATTAGCTTTGCATGCAATTTCGAAGCAATTCAGATGCCTAAAGGATG 1171
              | ||||||||     |||| ||| |  | || || |  || ||| | ||| | || || |
Db        929 AGTCATACACGTCGCTAGCATTGAAAACCATATCAAGACAGTTCCGTTGCTTGAAAGAGG 988

Qy       1172 CAATTGCTGAGCAAGTAAAGGCGACGAGCAAGAGTTTAGGTGAAGAGGAAGGCTTG---- 1227
              | || ||||  ||  |||| ||| | | ||||||| | || || || ||      ||
Db        989 CGATCGCTGGTCAGATAAAAGCGGCCAACAAGAGTCTTGGGGAGGAAGATTCAGTGTCTG 1048

Qy       1228 --GGAGGGAAAATCGAAGGCTCAAGACTCAAATTTGTGGACCATCATCTAAGGCAACAAC 1285
                 |  ||||  | || || || || ||||| || |||||||| ||  | || || |||
Db       1049 GTGTTGGGAGGTTTGAGGGGTCGAGGCTCAAGTTCGTGGACCACCACTTGAGACAGCAAA 1108

Qy       1286 GCGCGCTGCAACAGATAGGAATGATGCAAC---------CAAATGCTTGGAGACCCCAAA 1336
              | || || |||||  | |||||||| ||||          ||||||||||||| |||
Db       1109 GAGCTCTTCAACAACTGGGAATGATTCAACATCCTTCCAATAATGCTTGGAGACCTCAAC 1168

Qy       1337 GAGGTTTACCTGAAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTTTTCGAGCATTTTCTTC 1396
              | ||| | || ||| |||| ||||| || || |||||||||| ||||| || |||||||
Db       1169 GTGGTCTCCCAGAACGAGCCGTCTCAGTTCTCCGTGCTTGGCTCTTCGAACACTTTCTTC 1228

Qy       1397 ATCCTTACCCAAAGGATTCAGACAAAATCATGCTTGCTAAGCAAACGGGGCTAACAAGGA 1456
              |||| ||||| |||||||| |||||   |||||| |||||||||| || || ||  | |
Db       1229 ATCCATACCCTAAGGATTCGGACAAGCACATGCTAGCTAAGCAAACAGGACTCACTCGTA 1288

Qy       1457 GCCAGGTGTCTAACTGGTTCATAAATGCTCGAGTTCGATTATGGAAGCCAATGGTAGAAG 1516
              |||||||||| |||||||| ||||| ||  ||||||| |||||||| |||||||| || |
Db       1289 GCCAGGTGTCGAACTGGTTTATAAACGCGAGAGTTCGGTTATGGAAACCAATGGTGGAGG 1348

Qy       1517 AAATGTACTTGGAAGAAGTGAAGAATCAAGAACAAAACA 1555
              | |||||| |||| ||| ||||| | || | | | ||||
Db       1349 AGATGTACATGGAGGAAATGAAGGAGCAGGCAAAGAACA 1387


RESULT 10
US-10-225-066A-335
; Sequence 335, Application US/10225066A
; Publication No. US20030226173A1
; GENERAL INFORMATION:
;  APPLICANT: Mendel Biotechnology, Inc.
;  APPLICANT:  RATCLIFFE, Oliver
;  APPLICANT:  RIECHMANN, Jose Luis
;  APPLICANT:  ADAM, Luc J
;  APPLICANT:  DUBELL, Arnold T
;  APPLICANT:  HEARD, Jacqueline E
;  APPLICANT:  PILGRIM, Marsha L
;  APPLICANT:  JIANG, Cai-Zhong
;  APPLICANT:  REUBER, T. Lynne
;  APPLICANT:  CREELMAN, Robert A
;  APPLICANT:  PINEDA, Omaira
;  APPLICANT:  YU, Guo-Liang
;  APPLICANT:  BROUN, Pierre E
;  TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants
;  FILE REFERENCE: MBI0036-2 US
;  CURRENT APPLICATION NUMBER: US/10/225,066A
;  CURRENT FILING DATE:  2002-08-09
;  PRIOR APPLICATION NUMBER: 09/837,444
;  PRIOR FILING DATE: 2001-04-18
;  PRIOR APPLICATION NUMBER: 60/310,847
;  PRIOR FILING DATE: 2001-08-09
;  PRIOR APPLICATION NUMBER: 60/336,049
;  PRIOR FILING DATE: 2001-12-05
;  PRIOR APPLICATION NUMBER: 60/338,692
;  PRIOR FILING DATE: 2001-12-11
;  PRIOR APPLICATION NUMBER: 10/171,468
;  PRIOR FILING DATE: 2002-06-14
;  NUMBER OF SEQ ID NOS: 1122
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 335
;   LENGTH: 2385
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-10-225-066A-335

  Query Match            11.9%;  Score 324.6;  DB 17;  Length 2385;
  Best Local Similarity   65.2%;  Pred. No. 5.6e-65;
  Matches  534;  Conservative    0;  Mismatches  264;  Indels   21;  Gaps    3;

Qy        758 TTTTAGGCTCTAAGTATCTGAAAGCTGCACAAGAGCTTCTTGATGAAGTTGTTAATATTG 817
              || |  |||| |||||  |||| || ||||||||||||||||||| ||||| || ||   ||
Db        747 TTGTTAGCTCCAAGTACTTGAAGGCAGCACAAGAGCTTCTTGACGAAGTAGTCAACGCTG 806

Qy        818 TTGGAAAAAGCATCAAAGGAGATGATCAAAAGAAGGATAATTCAATGAATAAAGAATCAA 877
```

```
Db        807 ATTCCGATGACATGAACGCTAAATCCCAACTATTCTCATCGAAAAAGGGTAGTTGCGGAA 866

Qy        878 TGCCTTTGGCTAGTGATGTCAACACTAATAGTTCTGGTGGTGGTGAAAGTAGCAGCAGGC 937

Db        867 ATGATAAACCTGTCGGAGAATCATCGGCCGGCGCTGGAGGAGAAGGTTCCGGTGGCGGAG 926

Qy        938 AGAAAAATG------AAGTTGCTGTTGAGCTTACAACTGCTCAAAGACAAGAACTTCAAA 991

Db        927 CAGAAGCAGCCGGGAAACGTCCGGTGGAGCTAGGCACGGCAGAGAGACAAGAAATACAGA 986

Qy        992 TGAAAAAAGCCAAGCTTCTTGCCATGCTTGAAGAGGTGGAGCAAAGGTACAGACAGTACC 1051

Db        987 TGAAGAAAGCAAAACTTAGTAACATGCTTCATGAGGTGGAGCAGAGATATAGACAGTACC 1046

Qy       1052 ATCACCAAATGCAAATAATTGTATTATCATTTGAGCAAGTAGCAGGAATTGGATCAGCCA 1111

Db       1047 ACCAGCAGATGCAGATGGTGATCTCTTCGTTCGAGCAAGCGGCAGGGATAGGATCAGCGA 1106

Qy       1112 AATCATACACTCAATTAGCTTTGCATGCAATTTCGAAGCAATTCAGATGCCTAAAGGATG 1171

Db       1107 AGTCATACACGTCGCTAGCATTGAAAACCATATCAAGACAGTTCCGTTGCTTGAAAGAGG 1166

Qy       1172 CAATTGCTGAGCAAGTAAAGGCGACGAGCAAGAGTTTAGGTGAAGAGGAAGGCTTG---- 1227

Db       1167 CGATCGCTGGTCAGATAAAAGCGGCCAACAAGAGTCTTGGGGAGGAAGATTCAGTGTCTG 1226

Qy       1228 --GGAGGGAAAATCGAAGGCTCAAGACTCAAATTTGTGGACCATCATCTAAGGCAACAAC 1285

Db       1227 GTGTTGGGAGGTTTGAGGGGTCGAGGCTCAAGTTCGTGGACCACCACTTGAGACAGCAAA 1286

Qy       1286 GCGCGCTGCAACAGATAGGAATGATGCAAC---------CAAATGCTTGGAGACCCCAAA 1336

Db       1287 GAGCTCTTCAACAACTGGGAATGATTCAACATCCTTCCAATAATGCTTGGAGACCTCAAC 1346

Qy       1337 GAGGTTTACCTGAAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTTTTCGAGCATTTTCTTC 1396

Db       1347 GTGGTCTCCCAGAACGAGCCGTCTCAGTTCTCCGTGCTTGGCTCTTCGAACACTTTCTTC 1406

Qy       1397 ATCCTTACCCAAAGGATTCAGACAAAATCATGCTTGCTAAGCAAACGGGGCTAACAAGGA 1456

Db       1407 ATCCATACCCTAAGGATTCGGACAAGCACATGCTAGCTAAGCAAACAGGACTCACTCGTA 1466

Qy       1457 GCCAGGTGTCTAACTGGTTCATAAATGCTCGAGTTCGATTATGGAAGCCAATGGTAGAAG 1516

Db       1467 GCCAGGTGTCGAACTGGTTTATAAACGCGAGAGTTCGGTTATGGAAACCAATGGTGGAGG 1526

Qy       1517 AAATGTACTTGGAAGAAGTGAAGAATCAAGAACAAAACA 1555

Db       1527 AGATGTACATGGAGGAAATGAAGGAGCAGGCAAAGAACA 1565


RESULT 11
US-10-374-780A-2671
; Sequence 2671, Application US/10374780A
; Publication No. US20040019927A1
; GENERAL INFORMATION:
;  APPLICANT: Sherman, Bradley K
;  APPLICANT:  Riechmann, Jose Luis
;  APPLICANT:  Jiang, Cai-Zhong
;  APPLICANT:  Heard, Jacqueline E
;  APPLICANT:  Haake, Volker
;  APPLICANT:  Creelman, Robert A
;  APPLICANT:  Ratcliffe, Oliver
;  APPLICANT:  Adam, Luc J
;  APPLICANT:  Reuber, T. Lynne
;  APPLICANT:  Keddie, James
;  APPLICANT:  Broun, Pierre E
;  APPLICANT:  Pilgrim, Marsha L
;  APPLICANT:  Dubell III, Arnold T
;  APPLICANT:  Pineda, Omaira
;  APPLICANT:  Yu, Guo-Liang
;  TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
;  FILE REFERENCE: MBI-0047 CIP
;  CURRENT APPLICATION NUMBER: US/10/374,780A
;  CURRENT FILING DATE:  2003-02-25
;  PRIOR APPLICATION NUMBER: 09/837,944
;  PRIOR FILING DATE: 2001-04-18
;  PRIOR APPLICATION NUMBER: 60/310,847
;  PRIOR FILING DATE: 2001-08-09
;  PRIOR APPLICATION NUMBER: 09/934,455
;  PRIOR FILING DATE: 2001-08-22
;  PRIOR APPLICATION NUMBER: 60/336,049
;  PRIOR FILING DATE: 2001-11-19
;  PRIOR APPLICATION NUMBER: 60/338,692
;  PRIOR FILING DATE: 2001-12-11
;  PRIOR APPLICATION NUMBER: 10/171,468
;  PRIOR FILING DATE: 2002-06-14
;  PRIOR APPLICATION NUMBER: 10/225,066
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: 10/225,067
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: 10/225,068
;  PRIOR FILING DATE: 2002-08-09
;  NUMBER OF SEQ ID NOS: 2906
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2671
;   LENGTH: 2385
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   OTHER INFORMATION: G1589
US-10-374-780A-2671

  Query Match           11.9%;  Score 324.6;  DB 17;  Length 2385;
  Best Local Similarity   65.2%;  Pred. No. 5.6e-65;
  Matches  534;  Conservative    0;  Mismatches  264;  Indels   21;  Gaps    3;

Qy        758 TTTTAGGCTCTAAGTATCTGAAAGCTGCACAAGAGCTTCTTGATGAAGTTGTTAATATTG 817

Db        747 TTGTTAGCTCCAAGTACTTGAAGGCAGCACAAGAGCTTCTTGACGAAGTAGTCAACGCTG 806

Qy        818 TTGGAAAAAGCATCAAAGGAGATGATCAAAAGAAGGATAATTCAATGAATAAAGAATCAA 877

Db        807 ATTCCGATGACATGAACGCTAAATCCCAACTATTCTCATCGAAAAAGGGTAGTTGCGGAA 866

Qy        878 TGCCTTTGGCTAGTGATGTCAACACTAATAGTTCTGGTGGTGGTGAAAGTAGCAGCAGGC 937

Db        867 ATGATAAACCTGTCGGAGAATCATCGGCCGGCGCTGGAGGAGAAGGTTCCGGTGGCGGAG 926

Qy        938 AGAAAAATG------AAGTTGCTGTTGAGCTTACAACTGCTCAAAGACAAGAACTTCAAA 991

Db        927 CAGAAGCAGCCGGGAAACGTCCGGTGGAGCTAGGCACGGCAGAGAGACAAGAAATACAGA 986

Qy        992 TGAAAAAAGCCAAGCTTCTTGCCATGCTTGAAGAGGTGGAGCAAAGGTACAGACAGTACC 1051

Db        987 TGAAGAAAGCAAAACTTAGTAACATGCTTCATGAGGTGGAGCAGAGATATAGACAGTACC 1046

Qy       1052 ATCACCAAATGCAAATAATTGTATTATCATTTGAGCAAGTAGCAGGAATTGGATCAGCCA 1111

Db       1047 ACCAGCAGATGCAGATGGTGATCTCTTCGTTCGAGCAAGCGGCAGGGATAGGATCAGCGA 1106

Qy       1112 AATCATACACTCAATTAGCTTTGCATGCAATTTCGAAGCAATTCAGATGCCTAAAGGATG 1171

Db       1107 AGTCATACACGTCGCTAGCATTGAAAACCATATCAAGACAGTTCCGTTGCTTGAAAGAGG 1166

Qy       1172 CAATTGCTGAGCAAGTAAAGGCGACGAGCAAGAGTTTAGGTGAAGAGGAAGGCTTG---- 1227

Db       1167 CGATCGCTGGTCAGATAAAAGCGGCCAACAAGAGTCTTGGGGAGGAAGATTCAGTGTCTG 1226

Qy       1228 --GGAGGGAAAATCGAAGGCTCAAGACTCAAATTTGTGGACCATCATCTAAGGCAACAAC 1285

Db       1227 GTGTTGGGAGGTTTGAGGGGTCGAGGCTCAAGTTCGTGGACCACCACTTGAGACAGCAAA 1286

Qy       1286 GCGCGCTGCAACAGATAGGAATGATGCAAC---------CAAATGCTTGGAGACCCCAAA 1336
```

```
Db       1287 GAGCTCTTCAACAACTGGGAATGATTCAACATCCTTCCAATAATGCTTGGAGACCTCAAC 1346

Qy       1337 GAGGTTTACCTGAAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTTTTCGAGCATTTTCTTC 1396
              | ||| | || ||| |||| ||||| || || |||||||||||| ||||| || |||||||
Db       1347 GTGGTCTCCCAGAACGAGCCGTCTCAGTTCTCCGTGCTTGGCTCTTCGAACACTTTCTTC 1406

Qy       1397 ATCCTTACCCAAAGGATTCAGACAAAATCATGCTTGCTAAGCAAACGGGGCTAACAAGGA 1456
              |||| ||||| |||||||| |||||   |||||| |||||||||||| || || ||  | |
Db       1407 ATCCATACCCTAAGGATTCGGACAAGCACATGCTAGCTAAGCAAACAGGACTCACTCGTA 1466

Qy       1457 GCCAGGTGTCTAACTGGTTCATAAATGCTCGAGTTCGATTATGGAAGCCAATGGTAGAAG 1516
              |||||||||| |||||||| ||||| ||  ||||||| |||||||| |||||||| || |
Db       1467 GCCAGGTGTCGAACTGGTTTATAAACGCGAGAGTTCGGTTATGGAAACCAATGGTGGAGG 1526

Qy       1517 AAATGTACTTGGAAGAAGTGAAGAATCAAGAACAAAACA 1555
              | |||||| |||| ||| ||||| | || | | | ||||
Db       1527 AGATGTACATGGAGGAAATGAAGGAGCAGGCAAAGAACA 1565


RESULT 12
US-10-225-066A-335
; Sequence 335, Application US/10225066A
; Publication No. US20050160493A9
; GENERAL INFORMATION:
;  APPLICANT: Mendel Biotechnology, Inc.
;  APPLICANT:  RATCLIFFE, Oliver
;  APPLICANT:  RIECHMANN, Jose Luis
;  APPLICANT:  ADAM, Luc J
;  APPLICANT:  DUBELL, Arnold T
;  APPLICANT:  HEARD, Jacqueline E
;  APPLICANT:  PILGRIM, Marsha L
;  APPLICANT:  JIANG, Cai-Zhong
;  APPLICANT:  REUBER, T. Lynne
;  APPLICANT:  CREELMAN, Robert A
;  APPLICANT:  PINEDA, Omaira
;  APPLICANT:  YU, Guo-Liang
;  APPLICANT:  BROUN, Pierre E
;  TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants
;  FILE REFERENCE: MBI0036-2 US
;  CURRENT APPLICATION NUMBER: US/10/225,066A
;  CURRENT FILING DATE:  2002-08-09
;  PRIOR APPLICATION NUMBER: 09/837,444
;  PRIOR FILING DATE: 2001-04-18
;  PRIOR APPLICATION NUMBER: 60/310,847
;  PRIOR FILING DATE: 2001-08-09
;  PRIOR APPLICATION NUMBER: 60/336,049
;  PRIOR FILING DATE: 2001-12-05
;  PRIOR APPLICATION NUMBER: 60/338,692
;  PRIOR FILING DATE: 2001-12-11
;  PRIOR APPLICATION NUMBER: 10/171,468
;  PRIOR FILING DATE: 2002-06-14
;  NUMBER OF SEQ ID NOS: 1122
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 335
;   LENGTH: 2385
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-10-225-066A-335

  Query Match             11.9%;  Score 324.6;  DB 22;  Length 2385;
  Best Local Similarity   65.2%;  Pred. No. 5.6e-65;
  Matches  534;  Conservative    0;  Mismatches  264;  Indels   21;  Gaps    3;

Qy        758 TTTTAGGCTCTAAGTATCTGAAAGCTGCACAAGAGCTTCTTGATGAAGTTGTTAATATTG 817
              || |  |||| |||||  |||| || ||||||||||||||||| ||||| || ||   ||
Db        747 TTGTTAGCTCCAAGTACTTGAAGGCAGCACAAGAGCTTCTTGACGAAGTAGTCAACGCTG 806

Qy        818 TTGGAAAAAGCATCAAAGGAGATGATCAAAAGAAGGATAATTCAATGAATAAAGAATCAA 877
               |    |   ||| || |   |    |||                || |  ||       ||
Db        807 ATTCCGATGACATGAACGCTAAATCCCAACTATTCTCATCGAAAAAGGGTAGTTGCGGAA 866

Qy        878 TGCCTTTGGCTAGTGATGTCAACACTAATAGTTCTGGTGGTGGTGAAAGTAGCAGCAGGC 937
                  |    ||   | |      |     | |||| || |  |      |  || |
```

```
Db        867 ATGATAAACCTGTCGGAGAATCATCGGCCGGCGCTGGAGGAGAAGGTTCCGGTGGCGGAG 926

Qy        938 AGAAAAATG------AAGTTGCTGTTGAGCTTACAACTGCTCAAAGACAAGAACTTCAAA 991
                ||   |      ||  | | || |||||     || ||  | ||||||||| | || |
Db        927 CAGAAGCAGCCGGGAAACGTCCGGTGGAGCTAGGCACGGCAGAGAGACAAGAAATACAGA 986

Qy        992 TGAAAAAAGCCAAGCTTCTTGCCATGCTTGAAGAGGTGGAGCAAAGGTACAGACAGTACC 1051
              |||| ||||| || |||  |  ||||||| | |||||||||||| || || ||||||||||
Db        987 TGAAGAAAGCAAAACTTAGTAACATGCTTCATGAGGTGGAGCAGAGATATAGACAGTACC 1046

Qy       1052 ATCACCAAATGCAAATAATTGTATTATCATTTGAGCAAGTAGCAGGAATTGGATCAGCCA 1111
              | || || ||||| ||  |  | |  || || |||||||  ||||| || ||||||||  |
Db       1047 ACCAGCAGATGCAGATGGTGATCTCTTCGTTCGAGCAAGCGGCAGGGATAGGATCAGCGA 1106

Qy       1112 AATCATACACTCAATTAGCTTTGCATGCAATTTCGAAGCAATTCAGATGCCTAAAGGATG 1171
              | ||||||||     |||| ||| |  | || || |  || ||| | ||| | || || |
Db       1107 AGTCATACACGTCGCTAGCATTGAAAACCATATCAAGACAGTTCCGTTGCTTGAAAGAGG 1166

Qy       1172 CAATTGCTGAGCAAGTAAAGGCGACGAGCAAGAGTTTAGGTGAAGAGGAAGGCTTG---- 1227
              | || ||||  ||  |||| ||| | | ||||||| | || || || ||     ||
Db       1167 CGATCGCTGGTCAGATAAAAGCGGCCAACAAGAGTCTTGGGGAGGAAGATTCAGTGTCTG 1226

Qy       1228 --GGAGGGAAAATCGAAGGCTCAAGACTCAAATTTGTGGACCATCATCTAAGGCAACAAC 1285
                |  ||||   | || || || || ||||| || |||||||| ||  | || || |||
Db       1227 GTGTTGGGAGGTTTGAGGGGTCGAGGCTCAAGTTCGTGGACCACCACTTGAGACAGCAAA 1286

Qy       1286 GCGCGCTGCAACAGATAGGAATGATGCAAC---------CAAATGCTTGGAGACCCCAAA 1336
              | || || |||||  | |||||||| ||||          ||||||||||||| |||
Db       1287 GAGCTCTTCAACAACTGGGAATGATTCAACATCCTTCCAATAATGCTTGGAGACCTCAAC 1346

Qy       1337 GAGGTTTACCTGAAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTTTTCGAGCATTTTCTTC 1396
              | ||| | || ||| |||| ||||| || || |||||||||||| ||||| || |||||||
Db       1347 GTGGTCTCCCAGAACGAGCCGTCTCAGTTCTCCGTGCTTGGCTCTTCGAACACTTTCTTC 1406

Qy       1397 ATCCTTACCCAAAGGATTCAGACAAAATCATGCTTGCTAAGCAAACGGGGCTAACAAGGA 1456
              |||| ||||| |||||||| |||||   |||||| |||||||||||| || || ||  | |
Db       1407 ATCCATACCCTAAGGATTCGGACAAGCACATGCTAGCTAAGCAAACAGGACTCACTCGTA 1466

Qy       1457 GCCAGGTGTCTAACTGGTTCATAAATGCTCGAGTTCGATTATGGAAGCCAATGGTAGAAG 1516
              |||||||||| |||||||| ||||| ||  ||||||| |||||||| |||||||| || |
Db       1467 GCCAGGTGTCGAACTGGTTTATAAACGCGAGAGTTCGGTTATGGAAACCAATGGTGGAGG 1526

Qy       1517 AAATGTACTTGGAAGAAGTGAAGAATCAAGAACAAAACA 1555
              | |||||| |||| ||| ||||| | || | | | ||||
Db       1527 AGATGTACATGGAGGAAATGAAGGAGCAGGCAAAGAACA 1565


RESULT 13
US-10-425-114-9882
; Sequence 9882, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 9882
;   LENGTH: 1716
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: 700875318_FLI
US-10-425-114-9882

  Query Match             11.5%;  Score 314;  DB 18;  Length 1716;
```

```
  Best Local Similarity   77.6%;  Pred. No. 1.4e-62;
  Matches  380;  Conservative    0;  Mismatches  110;  Indels    0;  Gaps    0;

Qy       1062 GCAAATAATTGTATTATCATTTGAGCAAGTAGCAGGAATTGGATCAGCCAAATCATACAC 1121
              ||||||  | |||| ||||||||||||||| ||| ||   |||  | || ||||| |||||
Db          1 GCAAATTGTGGTATCATCATTTGAGCAAGCAGCGGGTTATGGGGCGGCAAAATCTTACAC 60

Qy       1122 TCAATTAGCTTTGCATGCAATTTCGAAGCAATTCAGATGCCTAAAGGATGCAATTGCTGA 1181
              |    | ||  |  |  | || || |||||||||| ||||| || ||||||||  |||
Db         61 TGCCCTTGCCCTAAAAACTATCTCAAAGCAATTCAGGTGCCTGAAAGATGCAATCTCTGC 120

Qy       1182 GCAAGTAAAGGCGACGAGCAAGAGTTTAGGTGAAGAGGAAGGCTTGGGAGGGAAAATCGA 1241
              ||| | || || |||||||||||  || ||||||||  ||  |||| ||||  ||| | ||
Db        121 ACAAATCAAAGCCACGAGCAAGACATTGGGTGAAGATGATTGCTTAGGAGTTAAAGTAGA 180

Qy       1242 AGGCTCAAGACTCAAATTTGTGGACCATCATCTAAGGCAACAACGCGCGCTGCAACAGAT 1301
              ||| ||||| || |  | ||| ||||| |||||  |||||||| | || || || ||  |
Db        181 AGGGTCAAGGCTTAGGTATGTTGACCACCATCTTCGGCAACAAAGGGCACTTCAGCAACT 240

Qy       1302 AGGAATGATGCAACCAAATGCTTGGAGACCCCAAAGAGGTTTACCTGAAAGAGCTGTCTC 1361
              ||||||||| || || ||||| ||||| ||||| ||||| || |||||| |||||||  ||
Db        241 AGGAATGATTCAGCCCAATGCATGGAGGCCCCAGAGAGGCTTGCCTGAACGAGCTGTTTC 300

Qy       1362 TGTCCTTCGTGCTTGGCTTTTCGAGCATTTTCTTCATCCTTACCCAAAGGATTCAGACAA 1421
                | ||||| |||||||||||| |||||||| ||||| || || |||||||| || ||  ||
Db        301 CATTCTTCGGGCTTGGCTTTTTGAGCATTTCCTTCACCCATATCCAAAGGACTCCGATAA 360

Qy       1422 AATCATGCTTGCTAAGCAAACGGGGCTAACAAGGAGCCAGGTGTCTAACTGGTTCATAAA 1481
              | | |||||||||||| ||||| || ||  | |||||||||||||| |||||||| || ||
Db        361 AGTTATGCTTGCTAAACAAACTGGACTTTCTAGGAGCCAGGTGTCAAACTGGTTTATCAA 420

Qy       1482 TGCTCGAGTTCGATTATGGAAGCCAATGGTAGAAGAAATGTACTTGGAAGAAGTGAAGAA 1541
              ||| ||||||||  | |||||||||||||| |||||||||||||||||||| | ||| |
Db        421 TGCCCGAGTTCGGCTTTGGAAGCCAATGGTTGAAGAAATGTACTTGGAAGAAATCAAGGA 480

Qy       1542 TCAAGAACAA 1551
              || ||||||
Db        481 ACATGAACAA 490
```

RESULT 14
US-10-424-599-7028
; Sequence 7028, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 7028
; LENGTH: 582
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_106354C.1
US-10-424-599-7028

```
  Query Match            11.1%;  Score 304;  DB 18;  Length 582;
  Best Local Similarity   76.4%;  Pred. No. 1.7e-60;
  Matches  373;  Conservative    0;  Mismatches  115;  Indels    0;  Gaps    0;

Qy        957 TGAGCTTACAACTGCTCAAAGACAAGAACTTCAAATGAAAAAAGCCAAGCTTCTTGCCAT 1016
              |||  |  ||||| |  | |||||||||| |||| ||||| |||||||| ||  | | |||
Db         95 TGAATTGTCAACTACGGAGAGACAAGAAATTCAGATGAAGAAAGCCAAACTAATCGACAT 154

Qy       1017 GCTTGAAGAGGTGGAGCAAAGGTACAGACAGTACCATCACCAAATGCAAATAATTGTATT 1076
              |||||| |||||||||||||||||||||| |||||||| || || ||||| ||| |  | |
Db        155 GCTTGATGAGGTGGAGCAAAGGTACAGGCAGTACCACCAGCAGATGCAGATAGTGATTTC 214

Qy       1077 ATCATTTGAGCAAGTAGCAGGAATTGGATCAGCCAAATCATACACTCAATTAGCTTTGCA 1136
              || || |||||||  || || || || ||||| |    | ||||||     |||| |||||
Db        215 CTCTTTCGAGCAAGCTGCGGGGATAGGCTCAGCGAGGACGTACACTGCCCTAGCATTGCA 274

Qy       1137 TGCAATTTCGAAGCAATTCAGATGCCTAAAGGATGCAATTGCTGAGCAAGTAAAGGCGAC 1196
              ||||||| |||||||| | ||  | |||||||||||||| | |||| |    ||  |
Db        275 AACAATTTCAAAGCAATTTCGGTGTTTGAAGGATGCAATTGCGGGGCAAATTCGAGCTGC 334

Qy       1197 GAGCAAGAGTTTAGGTGAAGAGGAAGGCTTGGGAGGGAAAATCGAAGGCTCAAGACTCAA 1256
              | ||||||||| || ||||||||  |||| |||||||||||| ||||| || || |||||
Db        335 AAACAAGAGTTTGGGGGAAGAGGATTGCTTTGGAGGGAAAATGGAAGGTTCTAGGCTCAA 394

Qy       1257 ATTTGTGGACCATCATCTAAGGCAACAACGCGCGCTGCAACAGATAGGAATGATGCAACCC 1316
              || ||| || |||||||  ||||||| || || || |||||  | ||||||| || |
Db        395 ATATGTTGATCATCATCTTCGGCAACAGCGAGCACTTCAACAATTGGGAATGATCCATCA 454

Qy       1317 AAATGCTTGGAGACCCCAAAGAGGTTTACCTGAAAGAGCTGTCTCTGTCCTTCGTGCTTG 1376
              |||||||||||||||| ||||| || ||||| ||| | || ||||| || ||||||||
Db        455 CAATGCTTGGAGACCCCAGAGAGGATTGCCTGAGAGATCGGTTTCTGTTCTGCGTGCTTG 514

Qy       1377 GCTTTTCGAGCATTTTCTTCATCCTTACCCAAAGGATTCAGACAAAATCATGCTTGCTAA 1436
              ||| || || || ||||| || ||||| || ||||| || ||||||  |||||||||  ||
Db        515 GCTCTTTGAACACTTTCTCCACCCTTATCCCAAGGACTCGGACAAACACATGCTTGCAAA 574

Qy       1437 GCAAACGG 1444
              ||||| |
Db        575 ACAAACAG 582
```

RESULT 15
US-10-437-963-41007
; Sequence 41007, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 41007
; LENGTH: 3028
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_44394C.1
US-10-437-963-41007

```
  Query Match            10.8%;  Score 295.6;  DB 19;  Length 3028;
  Best Local Similarity   60.3%;  Pred. No. 3.9e-58;
  Matches  528;  Conservative    0;  Mismatches  339;  Indels    9;  Gaps    2;

Qy        765 CTCTAAGTATCTGAAAGCTGCACAAGAGCTTCTTGATGAAGTTGTTAATATTGTTGGAAA 824
              ||| | |||||| || || |||||||| || |||||||| || || | | ||      |
Db       1135 CTCAAGGTATCTCAAGGCAGCACAAGAACTGCTTGATGAGGTCGTGAGTGTTTGGAAGAG 1194

Qy        825 AAGCATCAAAGGAGATGATCAAAAGAAGGATAATTCAATGAATAAAGAATCAATGCCTTT 884
              |  |   ||   | | |  || |||||| | | |||   ||   |||    |
Db       1195 CATAAAACAAAAGGCTCAGAAAGAGAAGGTTGAATCAGGAAAAGCAGATGGCAAA----- 1249

Qy        885 GGCTAGTGATGTCAACACTAATAGTTCTGGTGGTGGTGAAAGTAGCAGCAGGCAGAAAAA 944
              | | |||||      | || |||   ||||    |   | |
Db       1250 -GAGACTGATGGAGGGCCAAAAAGTGAGGGTGTATCTTCTAATCCACAAGAATCTGGCGC 1308
```

```
Qy        945 TGAAGTTGCTGTTGAGCTTACAACTGCTCAAAGACAAGAACTTCAAATGAAAAAAGCCAA 1004
              | | |  ||   ||||||| | |||||| | |  ||||| ||||| |  ||||  || ||
Db       1309 TAATGCAGCACCTGAGCTTTCCACTGCTGAGAAGCAAGAGCTTCAGAACAAAATGGCAAA 1368

Qy       1005 GCTTCTTGCCATGCTTGAAGAGGTGGAGCAAAGGTACAGACAGTACCATCACCAAATGCA 1064
               ||  | || ||| | || |||||||| |  |  |||| || ||  | ||||||||||
Db       1369 ACTGATGGCAATGTTGGATGAGGTGGACCGGAAATACAAGCATTATTACCACCAAATGCA 1428

Qy       1065 AATAATTGTATTATCATTTGAGCAAGTAGCAGGAATTGGATCAGCCAAATCATACACTCA 1124
              ||   | || | ||| |||||    ||||| |||   ||||| || ||  | |||||
Db       1429 AACCGTGGTTTCATCGTTTGATGTGGTAGCTGGACCAGGATCTGCGAAGCCTTACACAGC 1488

Qy       1125 ATTAGCTTTGCATGCAATTTCGAAGCAATTCAGATGCCTAAAGGATGCAATTGCTGAGCA 1184
              | | ||| | ||  |||| ||   ||| ||  | ||||| |||||||| ||   ||| ||
Db       1489 AGTTGCTCTTCAGACAATCTCACGGCACTTTCGGTGCCTGAAGGATGCTATCAATGATCA 1548

Qy       1185 AGTAAAGGCGACGAGCAAGAGTTTAGGTGAAGAGGAAGGCTTGGGAGGGAAAATCGAAGG 1244
                | || |  |  || || |   | || || || ||   ||     || ||    | |
Db       1549 GATCAATGTTATCAGGAAAAAGCTTGGCGAGGAAGAGAACTCATCTGGCAAGGAGGGAAA 1608

Qy       1245 CTCAA---GACTCAAATTTGTGGACCATCATCTAAGGCAACAACGCGCGCTGCAACAGAT 1301
              | ||   | |||   |   | || || || ||||||||||||||| ||  | ||||||
Db       1609 ATTAACCCGCCTCCGGTACATTGATCAGCAGCTAAGGCAACAACGTGCTTTCCAACAGTA 1668

Qy       1302 AGGAATGATGCAACCAAATGCTTGGAGACCCCAAAGAGGTTTACCTGAAAGAGCTGTCTC 1361
               || ||||| | || ||| || |||||||| || || ||  | |||||||   ||||  |
Db       1669 CGGTATGATTCCACAAAACGCCTGGAGACCACAGAGGGGACTGCCTGAAAACTCTGTTAC 1728

Qy       1362 TGTCCTTCGTGCTTGGCTTTTCGAGCATTTTCTTCATCCTTACCCAAAGGATTCAGACAA 1421
               | ||||| |||||||| || || || || || ||||| |||||||| ||||| || ||
Db       1729 AATTCTTCGCGCTTGGCTGTTTGAACACTTCCTCCATCCGTACCCAAAAGATTCCGAGAA 1788

Qy       1422 AATCATGCTTGCTAAGCAAACGGGGCTAACAAGGAGCCAGGTGTCTAACTGGTTCATAAA 1481
               | |||||||||| ||| || || || || ||||| ||| | || || ||||||||||
Db       1789 GTTAATGCTTGCTAGGCAGACTGGCCTGACTAGGAGTCAGATTTCGAATTGGTTCATAAA 1848

Qy       1482 TGCTCGAGTTCGATTATGGAAGCCAATGGTAGAAGAAATGTACTTGGAAGAAGTGAAGAA 1541
              ||| || || ||  | ||||| |||||| | ||||| ||||||   |||||  |     |
Db       1849 TGCCCGTGTCCGCCTTTGGAAACCAATGATTGAAGACATGTACAAAGAAGAGATTGGTGA 1908

Qy       1542 TCAAGAACAAAACAGTACTAATACTTCAGGAGATAACAAAAACAAAGAGACCAATATAAG 1601
              |  ||| ||| ||    |   | | || |   ||     ||  |   ||   || ||
Db       1909 TTTAGAGCAAGACTCCAACTCTTCCTCTGACAATGCACCAAGAAGTAAGGATAAAATGGC 1968

Qy       1602 TGCTCCAAATGAAGAGAAACATCCAATTATTACTAG 1637
                || || | ||  || |  |||  |  |   | ||
Db       1969 ATCTTCAGAAGACAAGGAGGATCTCAAAAACTCCAG 2004
```

Search completed: July 29, 2005, 14:48:36
Job time : 1683 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 29, 2005, 07:12:43 ; Search time 8585 Seconds
(without alignments)
12126.474 Million cell updates/sec

Title: US-10-624-201A-1
Perfect score: 2735
Sequence: 1 catgcagagataaaaatata.....gcaaaaaaaaaaaaaaaaa 2735

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	776.4	28.4	786	5	BQ509373	BQ509373 EST616788
c 2	722.2	26.4	743	5	BQ510344	BQ510344 EST617759
3	718.6	26.3	756	4	BG593861	BG593861 EST492539
4	666.4	24.4	699	5	BQ119770	BQ119770 EST605346
5	638.6	23.3	685	2	AW034617	AW034617 EST278301
6	608	22.2	708	4	BI920940	BI920940 EST540875
7	594.8	21.7	727	4	BI176426	BI176426 EST517371
8	581.6	21.3	644	2	AW930221	AW930221 EST340678
9	565	20.7	630	4	BG129304	BG129304 EST474950
10	564.2	20.6	789	4	BI921008	BI921008 EST540943
11	531	19.4	610	7	CO909070	CO909070 BJ02039F0
12	517.8	18.9	553	1	AI773253	AI773253 EST254353
c 13	502.4	18.4	721	2	BE435130	BE435130 EST406208
14	476	17.4	964	7	CK262102	CK262102 EST708180
15	467.8	17.1	495	2	BE460264	BE460264 EST411755
16	467.2	17.1	480	5	BQ509372	BQ509372 EST616787
17	446.2	16.3	623	2	AW441343	AW441343 EST310739
c 18	438	16.0	747	7	CV472924	CV472924 21081.1 D
19	427.8	15.6	479	2	BE433793	BE433793 EST404871
20	424.8	15.5	452	2	AW041164	AW041164 EST284028
21	419	15.3	752	4	BM535737	BM535737 EST588759
22	417.8	15.3	568	2	AW441354	AW441354 EST310750
23	417.6	15.3	440	2	AW093664	AW093664 EST286844
24	417.4	15.3	452	2	BF188184	BF188184 EST444471
25	403.8	14.8	576	1	AI894617	AI894617 EST264060
26	393.8	14.4	667	1	AI485358	AI485358 EST243679
27	392	14.3	740	7	CO111273	CO111273 GR_Eb004
28	391.4	14.3	554	1	AI897675	AI897675 EST267118
29	391.4	14.3	575	4	BI926891	BI926891 EST546780
30	389.6	14.2	494	2	AW094308	AW094308 EST287488
31	389.4	14.2	530	1	AI487525	AI487525 EST245847
32	388.6	14.2	523	1	AI898911	AI898911 EST268354
33	383.8	14.0	669	5	BQ148285	BQ148285 NF065F11F
34	383	14.0	783	2	AW688195	AW688195 NF004E06S
35	382.4	14.0	415	7	CV494999	CV494999 40158.1 C
36	377.2	13.8	645	4	BG887850	BG887850 EST513701
37	367.2	13.4	850	6	CB292855	CB292855 UCRCS01_0
38	366.8	13.4	1064	7	CV477853	CV477853 57798.1 D
39	363.6	13.3	844	7	CO101205	CO101205 GR_Eb002
40	361	13.2	636	4	BG887971	BG887971 EST513822
41	360.8	13.2	737	2	BE435482	BE435482 EST406560
42	356	13.0	529	4	BI930473	BI930473 EST550362
43	354.6	13.0	537	5	BM954132	BM954132 sam70e12.
44	352.4	12.9	636	2	AW442048	AW442048 EST311444
45	350.2	12.8	550	2	BE341111	BE341111 EST345206

ALIGNMENTS

RESULT 1
BQ509373/c
LOCUS BQ509373 786 bp mRNA linear EST 07-MAR-2003
DEFINITION EST616788 Generation of a set of potato cDNA clones for microarray analyses mixed potato tissues Solanum tuberosum cDNA clone STMHC24 3' end, mRNA sequence.
ACCESSION BQ509373
VERSION BQ509373.2 GI:21925113
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiids; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 786)
AUTHORS Buell,C.R., Hart,A., Baker,B., Tanksley,S., Fry,W., Smart,C., Restrepo,S., Griffiths,H., van der Hoeven,R., Tsai,J. and Karamycheva,S.A.
TITLE Generation of a set of potato cDNA clones for microarray analyses
JOURNAL Unpublished (2002)
COMMENT On Jun 10, 2002 this sequence version replaced gi:21368242.
Other_ESTs: EST616787
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
This clone can be obtained from the University of Arizona Genomics Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: T7.
FEATURES Location/Qualifiers
source 1. .786
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec or Binjte"
/db_xref="taxon:4113"
/clone="STMHC24"
/tissue_type="mixed tissues"
/lab_host="SOLR"
/clone_lib="Generation of a set of potato cDNA clones for microarray analyses mixed potato tissues"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; supplier: Combination of untreated and Phytophthora infestans-treated libraries of stolons, leaves, leaflets, axillary buds of stem explants, petioles, germinating eyes, tubers, or roots."
ORIGIN

Query Match 28.4%; Score 776.4; DB 5; Length 786;
Best Local Similarity 99.2%; Pred. No. 4.5e-172;
Matches 780; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy       1881 AGCTAGAGAATCATCAAATAAAGGGTTTACTAATCCTTTAATGGCAGCATACGCGATGGG 1940
             ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||
Db        786 AGCTAGAGAATCATCAAATAAAGGGTTTACTATCCCTTTAATGGCAGCATACGCGATGGA 727

Qy       1941 AGATTTTGGAAGGTTTGATCCTCATGATCAACAAATGACCGCGAATTTTCATGGAAATAA 2000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        726 GATTTTTGGAAGGTTTGATCCTCATGATCAACAAATGACCGCGAATTTTCATGGAAATAA 667

Qy       2001 TGGTGTCTCTCTTACTTTAGGACTTCCTCCTTCTGAAAACCTAGCCATGCCAGTGAGCCA 2060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        666 TGGTGTCTCTCTTACTTTAGGACTTCCTCCTTCTGAAAACCTAGCCATGCCAGTGAGCCA 607

Qy       2061 ACAAAATTACCTTTCTAATGACTTGGGAAGTAGGTCTGAAATGGGGAGTCATTACAATAG 2120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        606 ACAAAATTACCTTTCTAATGACTTGGGAAGTAGGTCTGAAATGGGGAGTCATTACAATAG 547

Qy       2121 AATGGGATATGAAAACATTGATTTTCAGAGTGGGAATAAGCGATTTCCGACTCAACTATT 2180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 AATGGGATATGAAAACATTGATTTTCAGAGTGGGAATAAGCGATTTCCGACTCAACTATT 487

Qy       2181 ACCAGATTTTGTTACAGGTAATCTAGGAACATGAATACCAGAAAGTCTCGTATTGATAGC 2240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        486 ACCAGATTTTGTTACAGGTAATCTAGGAACATGAATACCAGAAAGTCTCGTATTGATAGC 427

Qy       2241 TGAAAAGATAAAAGGAAGTTAGGGATACTCTTATATTGTGTGAGGCCTTCTGGCCCAAGT 2300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 TGAAAAGATAAAAGGAAGTTAGGGATACTCTTATATTGTGTGAGGCCTTCTGGCCCAAGT 367

Qy       2301 CGGAGGACCCAATTTGATACAACCTATCATAGGAGAAAAGAAGTGGAGACTAAATTAAAG 2360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 CGGAGGACCCAATTTGATACAACCTATCATAGGAGAAAAGAAGTGGAGACTAAATTAAAG 307

Qy       2361 TAACAAAATTTTAAAGCACACTTTCTAGTATATATACTTCTTTTTTTTATAGTATAGAAA 2420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 TAACAAAATTTTAAAGCACACTTTCTAGTATATATACTTCTTTTTTTTATAGTATAGAAA 247

Qy       2421 AGAAGAGATTTTGTGCTTTAGTGTATAGATAGAGTCTACTTAGTATAGGTTATACTTCTA 2480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 AGAAGAGATTTTGTGCTTTAGTGTATAGATAGAGTCTACTTAGTATAGGTTATACTTCTA 187

Qy       2481 GTTCCTTGAGAAGATTGATACAACTAGTAGTATTTTTTTTCTTTTGGGTTGGCTTGGAGT 2540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 GTTCCTTGAGAAGATTGATACAACTAGTAGTATTTTTTTTCTTTTGGGTTGGCTTGGAGT 127

Qy       2541 ACTATTTTAAGTTATTGGAAACTAGCTATAGTAAATGTTGTAAAGTTGTGATATTGTTCC 2600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 ACTATTTTAAGTTATTGGAAACTAGCTATAGTAAATGTTGTAAAGTTGTGATATTGTTCC 67

Qy       2601 TCTCAATTTGCATATAATTTGAAATATTTTGTACCTACTAGCTAGTCTCTAAATTATGTT 2660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          66 TCTCAATTTGCATATAATTTGAAATATTTTGTACCTACTAGCTAGTCTCTAAATTATGTT 7

Qy       2661 TCCATT 2666
             ||||||
Db           6 TCCATT 1
```

RESULT 2
BQ510344/c
LOCUS BQ510344 743 bp mRNA linear EST 07-MAR-2003
DEFINITION EST617759 Generation of a set of potato cDNA clones for microarray analyses mixed potato tissues Solanum tuberosum cDNA clone STMHJ58 3' end, mRNA sequence.
ACCESSION BQ510344
VERSION BQ510344.2 GI:21926041
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiids; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 743)
AUTHORS Buell,C.R., Hart,A., Baker,B., Tanksley,S., Fry,W., Smart,C., Restrepo,S., Griffiths,H., van der Hoeven,R., Tsai,J. and Karamycheva,S.A.
TITLE Generation of a set of potato cDNA clones for microarray analyses
JOURNAL Unpublished (2002)
COMMENT On Jun 10, 2002 this sequence version replaced gi:21369213.
Other_ESTs: EST617758
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
This clone can be obtained from the University of Arizona Genomics Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: T7.
FEATURES Location/Qualifiers
source 1. .743
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec or Binjte"
/db_xref="taxon:4113"
/clone="STMHJ58"
/tissue_type="mixed tissues"
/lab_host="SOLR"
/clone_lib="Generation of a set of potato cDNA clones for microarray analyses mixed potato tissues"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; supplier: Combination of untreated and Phytophthora infestans-treated libraries of stolons, leaves, leaflets, axillary buds of stem explants, petioles, germinating eyes, tubers, or roots."
ORIGIN

Query Match 26.4%; Score 722.2; DB 5; Length 743;
Best Local Similarity 99.5%; Pred. No. 2.8e-159;
Matches 735; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

```
Qy       1811 AAAAGCAAAGAAATGACATGCACAAGTTTTCTCCAAGTAGTATTCTTTCATCTGTTGACA 1870
             |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db        743 AAAAGCAAAGAAATGACATGCACAAGTTTTCTCCAAGTAGTATTCTTTCATCTGTGGACA 684

Qy       1871 TGGAAGCCAAAGCTAGAGAATCATCAAATAAAGGGTTTACTAATCCTTTAATGGCAGCAT 1930
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        683 TGGAAGCCAAAGCTAGAGAATCATCAAATAAAGGGTTTACTAATCCTTTAATGGCAGCAT 624

Qy       1931 ACGCGATGGGAGATTTTGGAAGGTTTGATCCTCATGATCAACAAATGACCGCGAATTTTC 1990
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        623 ACGCGATGGGAGATTTTGGAAGGTTTGATCCTCATGATCAACAAATGACCGCGAATTTTC 564

Qy       1991 ATGGAAATAATGGTGTCTCTCTTACTTTAGGACTTCCTCCTTCTGAAAACCTAGCCATGC 2050
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        563 ATGGAAATAATGGTGTCTCTCTTACTTTAGGACTTCCTCCTTCTGAAAACCTAGCCATGC 504

Qy       2051 CAGTGAGCCAACAAAATTACCTTTCTAATGACTTGGGAAGTAGGTCTGAAATGGGGAGTC 2110
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        503 CAGTGAGCCAACAAAATTACCTTTCTAATGACTTGGGAAGTAGGTCTGAAATGGGGAGTC 444

Qy       2111 ATTACAATAGAATGGGATATGAAAACATTGATTTTCAGAGTGGGAATAAGCGATTTCCGA 2170
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        443 ATTACAATAGAATGGGATATGAAAACATTGATTTTCAGAGTGGGAATAAGCGATTTCCGA 384

Qy       2171 CTCAACTATTACCAGATTTTGTTACAGGTAATCTAGGAACATGAATACCAGAAAGTCTCG 2230
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        383 CTCAACTATTACCAGATTTTGTTACAGGTAATCTAGGAACATGAATACCAGAAAGTCTCG 324

Qy       2231 TATTGATAGCTGAAAAGATAAAAGGAAGTTAGGGATACTCTTATATTGTGTGAGGCCTTC 2290
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 TATTGATAGCTGAAAAGATAAAAGGAAGTTAGGGATACTCTTATATTGTGTGAGGCCTTC 264
```

```
Qy       2291 TGGCCCAAGTCGGAGGACCCAATTTGATACAACCTATCATAGGAGAAAAGAAGTGGAGAC 2350
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 TGGCCCAAGTCGGAGGACCCAATTTGATACAACCTATCATAGGAGAAAAGAAGTGGAGAC 204

Qy       2351 TAAATTAAAGTAACAAAATTTTAAAGCACACTTTCTAGTATATATACTTCTTTTTTTTAT 2410
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 TAAATTAAAGTAACAAAATTTTAAAGCACACTTTCTAGTATATATACTTCTTTTTTTTAT 144

Qy       2411 AGTATAGAAAAGAAGAGATTTTGTGCTTTAGTGTATAGATAGAGTCTACTTAGTATAGGT 2470
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 AGTATAGAAAAGAAGAGATTTTGTGCTTTAGTGTATAGATAGAGTCTACTTAGTATAGGT 84

Qy       2471 TATACTTCTAGTTCCTTGAGAAGATTGATACAACTAGTAGTATTTTTTTTCTTTTGGGTT 2530
             |||||||||||||| |||||||||||||||||||||||| ||||||||||||||||||
Db          83 TATACTTCTAGTTCCCTGAGAAGATTGATACAACTAGTAGTA-TTTTTTTCTTTTGGGTT 25

Qy       2531 GGCTTGGAGTACTATTTTA 2549
             |||||||||||||||||| |
Db          24 GGCTTGGAGTACTATTTAA 6
```

RESULT 3
BG593861
LOCUS BG593861 756 bp mRNA linear EST 07-MAR-2003
DEFINITION EST492539 cSTS Solanum tuberosum cDNA clone cSTS5P9 5' sequence, mRNA sequence.
ACCESSION BG593861
VERSION BG593861.1 GI:13612001
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiids; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 756)
AUTHORS van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Chiemingo,A., Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.
TITLE Generations of ESTs from sprouting potato eyes
JOURNAL Unpublished (2000)
COMMENT Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
This clone can be obtained from the University of Arizona Genomics Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: M13F-R.
FEATURES Location/Qualifiers
source 1. .756
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSTS5P9"
/tissue_type="sprouting eyes from tubers"
/dev_stage="12-14 weeks post harvest"
/lab_host="SOLR"
/clone_lib="cSTS"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; Various sizes of sprouting eyes (2mm to 15mm) were taken from tubers. The tubers were incubated at 26C in the dark for 2-3 weeks prior to sprouting. The eyes were frozen in liquid nitrogen immediately upon removal from tubers."
ORIGIN

Query Match 26.3%; Score 718.6; DB 4; Length 756;
Best Local Similarity 98.1%; Pred. No. 2e-158;
Matches 741; Conservative 0; Mismatches 4; Indels 10; Gaps 1;

```
Qy         41 AACTTCTCAAAGCTTAGAGAGCTACCACCCGAAGATAGACAGTTAGTTACA--------- 91
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 AACTTCTCAAAGCTTAGAGAGCTACCACCCGAAGATAGACAGTTAGTTACATGTGTGTAT 61

Qy         92 -TGTACTGTTATAGATAAAAGGAGAAATCCGAAGAAGAAAGAATTTTTTTTGCAGATATG 150
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GTGTACTGTTATAGATAAAAGGAGAAATCCGAAGAAGAAAGAATTTTTTTTGCAGATATG 121

Qy        151 TACTATCAAGGAACCTCGGATAATACTAATATACAAGCTGATCATCAACAACGTCATAAT 210
              |||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        122 TACTATCAAGGAACCTCGGATAATACTAATATACAAGCTGATCATCAACAACATCATAAT 181

Qy        211 CATGGGAATAGTAATAATAATAATATTCAGACACTTTATTTGATGAACCCTAACAATTAT 270
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 CATGGGAATAGTAATAATAATAATATTCAGACACTTTATTTGATGAACCCTAACAATTAT 241

Qy        271 ATGCAAGGCTACACTACTTCTGACACACAGCAGCAGCAGCAGTTACTTTTCCTGAATTCT 330
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 ATGCAAGGCTACACTACTTCTGACACACAGCAGCAGCAGCAGTTACTTTTCCTGAATTCT 301

Qy        331 TCACCAGCAGCAAGCAACGCGCTTTGCCATGCGAATATACAACACGCGCCGCTGCAACAG 390
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 TCACCAGCAGCAAGCAACGCGCTTTGCCATGCGAATATACAACACGCGCCGCTGCAACAG 361

Qy        391 CAGCACTTTGTCGGTGTGCCTCTTCCGGCAGTAAGTTTGCACGATCAGATCAATCATCAT 450
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 CAGCACTTTGTCGGTGTGCCTCTTCCGGCAGTAAGTTTGCACGATCAGATCAATCATCAT 421

Qy        451 GGACTTTTACAGCGCATGTGGAACAACCAAGATCAATCTCAGCAGGTGATAGTACCATCG 510
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 GGACTTTTACAGCGCATGTGGAACAACCAAGATCAATCTCAGCAGGTGATAGTACCATCG 481

Qy        511 TCGACGGGGGTTTCTGCCACGTCATGTGGCGGGATCACCACGGACTTGGCGTCTCAATTG 570
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 TCGACGGGGGTTTCTGCCACGTCATGTGGCGGGATCACCACGGACTTGGCGTCTCAATTG 541

Qy        571 GCGTTTCAGAGGCCGATTCCGACACCACAACACCGACAGCAGCAACAACAGCAAGGCGGT 630
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 GCGTTTCAGAGGCCGATTCCGACACCACAACACCGACAGCAGCAACAACAGCAAGGCGGT 601

Qy        631 CTATCTCTAAGCCTTTCTCCTCAGCTACAACAGCAAATTAGTTTCAATAACAATATTTCA 690
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 CTATCTCTAAGCCTTTCTCCTCAGCTACAACAGCAAATTAGTTTCAATAACAATATTTCA 661

Qy        691 TCCTCATCACCAAGGACAAATAATGTTACTATTAGGGGAACATTAGATGGAAGTTCTAGC 750
              |||||||||||||||||||||||||||||||||||||||||||||||| | ||||||
Db        662 TCCTCATCACCAAGGACAAATAATGTTACTATTAGGGGAACATTAGATGGGAAGTCTAGC 721

Qy        751 AACATGGTTTTAGGCTCTAAGTATCTGAAAGCTGC 785
              |||||||||||||||||||||||||||||||||||
Db        722 AACATGGTTTTAGGCTCTAAGTATCTGAAAGCTGC 756
```

RESULT 4
BQ119770
LOCUS BQ119770 699 bp mRNA linear EST 07-MAR-2003
DEFINITION EST605346 mixed potato tissues Solanum tuberosum cDNA clone STMEM54 5' end, mRNA sequence.
ACCESSION BQ119770
VERSION BQ119770.2 GI:21919664
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiids; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 699)
AUTHORS Buell,C.R., Hart,A., Baker,B., Tanksley,S., Fry,W., Smart,C., Restrepo,S., Griffiths,H., van der Hoeven,R., Tsai,J. and Karamycheva,S.A.
TITLE Generation of a set of potato cDNA clones for microarray analyses
JOURNAL Unpublished (2002)
COMMENT On Apr 17, 2002 this sequence version replaced gi:20171732.
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA

```
            Email: potato-array@tigr.org
            This clone can be obtained from the University of Arizona Genomics
            Institute. Orders can be made through URL:
            http://genome.arizona.edu/orders/
            Seq primer: T3.
FEATURES              Location/Qualifiers
     source           1. .699
                      /organism="Solanum tuberosum"
                      /mol_type="mRNA"
                      /cultivar="Kennebec or Binjte"
                      /db_xref="taxon:4113"
                      /clone="STMEM54"
                      /tissue_type="mixed tissues"
                      /lab_host="SOLR"
                      /clone_lib="mixed potato tissues"
                      /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
                      XhoI; supplier: Combination of untreated and Phytophthora
                      infestans-treated libraries of stolons, leaves, leaflets,
                      axillary buds of stem explants, petioles, germinating
                      eyes, tubers, or roots."
ORIGIN

  Query Match             24.4%;  Score 666.4;  DB 5;  Length 699;
  Best Local Similarity   98.4%;  Pred. No. 4.1e-146;
  Matches  687;  Conservative    0;  Mismatches    1;  Indels   10;  Gaps    1;

Qy         12 AAAAATATAGATCAGTCTGACAAGAAGGCAACTTCTCAAAGCTTAGAGAGCTACCACCCG 71
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 AAAAATATAGATCAGTCTGACAAGAAGGCAACTTCTCAAAGCTTAGAGAGCTACCACCCG 61

Qy         72 AAGATAGACAGTTAGTTACA----------TGTACTGTTATAGATAAAAGGAGAAATCCG 121
              |||||||||||||||||||          ||||||||||||||||||||||||||||||
Db         62 AAGATAGACAGTTAGTTACATGTGTGTATGTGTACTGTTATAGATAAAAGGAGAAATCCG 121

Qy        122 AAGAAGAAAGAATTTTTTTTGCAGATATGTACTATCAAGGAACCTCGGATAATACTAATA 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 AAGAAGAAAGAATTTTTTTTGCAGATATGTACTATCAAGGAACCTCGGATAATACTAATA 181

Qy        182 TACAAGCTGATCATCAACAACGTCATAATCATGGGAATAGTAATAATAATAATATTCAGA 241
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        182 TACAAGCTGATCATCAACAACATCATAATCATGGGAATAGTAATAATAATAATATTCAGA 241

Qy        242 CACTTTATTTGATGAACCCTAACAATTATATGCAAGGCTACACTACTTCTGACACACAGC 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CACTTTATTTGATGAACCCTAACAATTATATGCAAGGCTACACTACTTCTGACACACAGC 301

Qy        302 AGCAGCAGCAGTTACTTTTCCTGAATTCTTCACCAGCAGCAAGCAACGCGCTTTGCCATG 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 AGCAGCAGCAGTTACTTTTCCTGAATTCTTCACCAGCAGCAAGCAACGCGCTTTGCCATG 361

Qy        362 CGAATATACAACACGCGCCGCTGCAACAGCAGCACTTTGTCGGTGTGCCTCTTCCGGCAG 421
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 CGAATATACAACACGCGCCGCTGCAACAGCAGCACTTTGTCGGTGTGCCTCTTCCGGCAG 421

Qy        422 TAAGTTTGCACGATCAGATCAATCATCATGGACTTTTACAGCGCATGTGGAACAACCAAG 481
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 TAAGTTTGCACGATCAGATCAATCATCATGGACTTTTACAGCGCATGTGGAACAACCAAG 481

Qy        482 ATCAATCTCAGCAGGTGATAGTACCATCGTCGACGGGGGTTTCTGCCACGTCATGTGGCG 541
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 ATCAATCTCAGCAGGTGATAGTACCATCGTCGACGGGGGTTTCTGCCACGTCATGTGGCG 541

Qy        542 GGATCACCACGGACTTGGCGTCTCAATTGGCGTTTCAGAGGCCGATTCCGACACCACAAC 601
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 GGATCACCACGGACTTGGCGTCTCAATTGGCGTTTCAGAGGCCGATTCCGACACCACAAC 601

Qy        602 ACCGACAGCAGCAACAACAGCAAGGCGGTCTATCTCTAAGCCTTTCTCCTCAGCTACAAC 661
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 ACCGACAGCAGCAACAACAGCAAGGCGGTCTATCTCTAAGCCTTTCTCCTCAGCTACAAC 661

Qy        662 AGCAAATTAGTTTCAATAACAATATTTCATCCTCATCA 699
              |||||||||||||||||||||||||||||||||||||
Db        662 AGCAAATTAGTTTCAATAACAATATTTCATCCTCATCA 699
```

```
RESULT 5
AW034617
LOCUS       AW034617                685 bp    mRNA    linear   EST 18-MAY-2001
DEFINITION  EST278301 tomato callus, TAMU Lycopersicon esculentum cDNA clone
            cLEC33H16, mRNA sequence.
ACCESSION   AW034617
VERSION     AW034617.1  GI:5893373
KEYWORDS    EST.
SOURCE      Lycopersicon esculentum (tomato)
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1  (bases 1 to 685)
  AUTHORS   Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T.,
            Holt,I.E., Liang,F., Upton,J., Craven,M.B., Bowman,C.L., Ahn,S.,
            Ronning,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and
            Giovannoni,J.
  TITLE     Generation of ESTs from tomato callus tissue
  JOURNAL   Unpublished (1999)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            5 prime sequence.
FEATURES              Location/Qualifiers
     source           1. .685
                      /organism="Lycopersicon esculentum"
                      /mol_type="mRNA"
                      /cultivar="TA496"
                      /db_xref="taxon:4081"
                      /clone="cLEC33H16"
                      /tissue_type="callus"
                      /dev_stage="25-40 days old"
                      /lab_host="XL1-Blue MRF'"
                      /clone_lib="tomato callus, TAMU"
                      /note="Vector: pBlueScript SK(-); Site_1: EcoR1; Site_2:
                      Xho1; supplier: Giovannoni laboratory; cLEC - Cotyledons
                      of seedlings 7-10 days post-germination were excised, cut
                      at both ends and placed on MS medium with no selection.
                      Mixed callus was harvested at 25 and 40 days and included
                      undifferentiated masses. Tomato Callus EST Library"
ORIGIN

  Query Match             23.3%;  Score 638.6;  DB 2;  Length 685;
  Best Local Similarity   96.4%;  Pred. No. 1.5e-139;
  Matches  666;  Conservative    0;  Mismatches   19;  Indels    6;  Gaps    1;

Qy       1108 GCCAAATCATACACTCAATTAGCTTTGCATGCAATTTCGAAGCAATTCAGATGCCTAAAG 1167
              || |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db          1 GCAAAATCATACACTCAATTAGCTTTGCATGCAATTTCAAAGCAATTCAGATGCCTAAAG 60

Qy       1168 GATGCAATTGCTGAGCAAGTAAAGGCGACGAGCAAGAGTTTAGGTGAAGAGGAAGGCTTG 1227
              |||||||||  ||||||||| |||||||||||||||||||||||||||| |||||||||
Db         61 GATGCAATTAGTGAGCAAGTGAAGGCGACGAGCAAGAGTTTAGGTGAAGATGAAGGCTTG 120

Qy       1228 GGAGGGAAAATCGAAGGCTCAAGACTCAAATTTGTGGACCATCATCTAAGGCAACAACGC 1287
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAGGGAAAATTGAAGGCTCAAGACTCAAATTTGTGGACCATCATCTAAGGCAACAACGC 180

Qy       1288 GCGCTGCAACAGATAGGAATGATGCAACCAAATGCTTGGAGACCCCAAAGAGGTTTACCT 1347
              |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCGCTGCAACAGCTAGGAATGATGCAACCAAATGCTTGGAGACCCCAAAGAGGTTTACCT 240

Qy       1348 GAAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTTTTCGAGCATTTTCTTCATCCTTACCCA 1407
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTTTTCGAGCATTTTCTTCATCCTTACCCC 300

Qy       1408 AAGGATTCAGACAAAATCATGCTTGCTAAGCAAACGGGGCTAACAAGGAGCCAGGTGTCT 1467
              || || |||||||||||||||||||||||||||||||||||||||||||||||||| |||
```

```
Db        301 AAAGACTCAGACAAAATCATGCTTGCTAAGCAAACGGGGCTAACAAGGAGCCAGGTCTCT 360

Qy       1468 AACTGGTTCATAAATGCTCGAGTTCGATTATGGAAGCCAATGGTAGAAGAAATGTACTTG 1527
              |||||||||| |||||||||||||||||||||||||||||| ||||||||||||||||
Db        361 AACTGGTTCATTAATGCTCGAGTTCGATTATGGAAGCCAATGGTTGAAGAAATGTACTTG 420

Qy       1528 GAAGAAGTGAAGAATCAAGAACAAAACAGTACTAATACTTCAGGAGATAACAAAAACAAA 1587
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db        421 GAAGAAGTGAAGAATCAAGAACAAAACAGTAGTAATACTTCAGGAGATAACAAAAACAAA 480

Qy       1588 GAGACCAATATAAGTGCTCCAAATGAAGAGAAACATCCAATTATTACTAGCAGCTTATTA 1647
              ||||| |||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db        481 GAGACGAATATAAGTGCTCCAAATGAAGAGAAACAACCAATTATTACTAGCAGCTTATTA 540

Qy       1648 CAAGATGGTATTACTACTACTCAAGCAGAAATTTCTACCTCAACTATTTCAACTTCCCCT 1707
              ||||||||      ||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CAAGATGG------TACTACTCAAGCAGAAATTTCTACCTCAACTATTTCAACTTCCCCT 594

Qy       1708 ACTGCAGGTGCTTCACTTCATCATGCTCACAATTTCTCCTTCCTTGGTTCATTCAACATG 1767
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        595 ACTGCAGGTGCTTCACTTCATCATGCTCACAACTTCTCCTTCCTTGGTTCATTCAACATG 654

Qy       1768 GATAATACTACTACTACTGTTGATCATATTG 1798
              || |||||||||||||||||||||||||||||
Db        655 GAGAATACTACTACTACTGTTGATCATATTG 685


RESULT 6
BI920940
LOCUS       BI920940                708 bp    mRNA    linear   EST 10-MAR-2003
DEFINITION  EST540875 potato microtubers, in vitro-grown Solanum tuberosum cDNA
            clone cSTE26N18 5' end, mRNA sequence.
ACCESSION   BI920940
VERSION     BI920940.1  GI:16216968
KEYWORDS    EST.
SOURCE      Solanum tuberosum (potato)
  ORGANISM  Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum.
REFERENCE   1  (bases 1 to 708)
  AUTHORS   van der Hoeven,R., Bezzerides,J., Bachem,C., Visser,R.,
            Karamycheva,S.A., Tsai,J., Van Aken,S., Utterback,T., Chiemingo,A.,
            Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.
  TITLE     Generation of ESTs from in vitro grown microtubers (2001b)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Robin Buell
            The Institute for Genomic Research
            9712 Medical Center Dr, Rockville, MD 20850, USA
            Email: potato-array@tigr.org
            This clone can be obtained from the University of Arizona Genomics
            Institute. Orders can be made through URL:
            http://genome.arizona.edu/orders/
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1. .708
                     /organism="Solanum tuberosum"
                     /mol_type="mRNA"
                     /cultivar="Bintje"
                     /db_xref="taxon:4113"
                     /clone="cSTE26N18"
                     /tissue_type="axillary buds of stem explants; growing
                     sink-tubers"
                     /dev_stage="7, 8 and 10 days"
                     /lab_host="SOLR"
                     /clone_lib="potato microtubers, in vitro-grown"
                     /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
                     XhoI; Tissue supplied by Christian Bachem and Richard
                     Visser (Department of Plant Breeding, Wageningen
                     University, The Netherlands); sequencing by The Institute
                     for Genomic Research. The cSTA libraries will attempt to
                     capture the induction and initiation/initial growth of the
                     tuber in an in vitro system as described in Bachem et al.
                     (Plant Journal, 1996). Small microtubers develop from
                     axillary buds attached to stem explants when placed on a
                     high sucrose medium (10%). Visible morphological changes
                     occur synchronously at day five in the axillary buds. The
                     first library, cSTA (1-20) consists of axillary buds
                     harvested on days 1-3. This targets those genes involved
                     in induction of the microtubers. The following libraries,
                     cSTA (21-40) and cSTA (41-60), capture genes involved in
                     tuber initiation and outgrowth. This library is noted as
                     P3  in Tanksley lab notebooks."
ORIGIN

  Query Match             22.2%;  Score 608;  DB 4;  Length 708;
  Best Local Similarity   94.2%;  Pred. No. 2.5e-132;
  Matches  669;  Conservative    0;  Mismatches   25;  Indels   16;  Gaps    3;

Qy         50 AAGCTTAGAGAGCTACCACCCGAAGATAGACAGTTAGTTACA----------TGTACTGT 99
              ||| ||||||||||||||||||||||||||||||||||||||          ||||||||
Db          2 AAGTTTAGAGAGCTACCACCCGAAGATAGACAGTTAGTTACATGTGTGTATGTGTACTGT 61

Qy        100 TATAGATAAAAGGAGAAATCCGAAGAAGAAAGAATTTTTTTTGCAGATATGTACTATCAA 159
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 TATAGATAAAAGGAGAAATCCGAAGAAGAAAGAATTTTTTTTGCAGATATGTACTATCAA 121

Qy        160 GGAACCTCGGATAATACTAATATACAAGCTGATCATCAACAACGTCATAATCATGGGAAT 219
              |||||||||||||   |||||||||||||||||||||||| ||||||||||||||||
Db        122 GGAACCTCGGATAA---TAATATACAAGCTGATCATCAACAACATCATAATCATGGGAAT 178

Qy        220 AGTAATAATAATAATATTCAGACACTTTATTTGATGAACCCTAACAATTATATGCAAGGC 279
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        179 AGTAATAATAATAATATTCAGACACTTTATTTGATGAACCCTAACAATTATATGCAAGGC 238

Qy        280 TACACTACTTCTGACACA---CAGCAGCAGCAGCAGTTACTTTTCCTGAATTCTTCACCA 336
              ||||||||||||||||||   |||||||||||||||||||||||||||||||||||||||
Db        239 TACACTACTTCTGACACACAGCAGCAGCAGCAGCAGTTACTTTTCCTGAATTCTTCACCA 298

Qy        337 GCAGCAAGCAACGCGCTTTGCCATGCGAATATACAACACGCGCCGCTGCAACAGCAGCAC 396
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 GCAGGAAGCAACGCGCTTTGCCATGCGAATATACAACACGCGCCGCTGCAACAGCAGCAC 358

Qy        397 TTTGTCGGTGTGCCTCTTCCGGCAGTAAGTTTGCACGATCAGATCAATCATCATGGACTT 456
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 TTTGTCGGTGTGCCTCTTCCGGCAGTAAGTTTGCACGATCAGATCAATCATCATGGACTT 418

Qy        457 TTACAGCGCATGTGGAACAACCAAGATCAATCTCAGCAGGTGATAGTACCATCGTCGACG 516
              |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        419 TTACAGCGTATGTGGAACAACCAAGATCAATCTCAGCAGGTGATAGTACCATCGTCGACG 478

Qy        517 GGGGTTTCTGCCACGTCATGTGGCGGGATCACCACGGACTTGGCGTCTCAATTGGCGTTT 576
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        479 GGGGTTTCTGCCACGTCATGTGGCGGGATTACCACGGACTTGGCGTCTCAATTGGCGTTT 538

Qy        577 CAGAGGCCGATTCCGACACCACAACACCGACAGCAGCAACAACAGCAAGGCGGTCTATCT 636
              |||||||||||  |    | ||||||||||||||||||||||||||||||||||||||
Db        539 CAGAGGCCGATTGTGGTGTCGGAACACCGACAGCAGCAACAACAGCAAGGCGGTCTATCT 598

Qy        637 CTAAGCCTTTCTCCTCAGCTACAACAGCAAATTAGTTTCAATAACAATATTTCATCCTCA 696
              |||||||||||||||||| |||||||||||||||||||||||||||||||| |||||
Db        599 CTAAGCCTTTCTCCTCAGCAACAACAGCAAATTAGTTTCAATAACAATATTTCAGCCTCA 658

Qy        697 TCACCAAGGACAAATAATGTTACTATTAGGGGAACATTAGATGGAAGTTC 746
              ||||| | |   ||||||||||||||||| |||| ||||||||| |  ||
Db        659 TCACCCAAGGACAATAATGTTACTATTAGAGGAAGATTAGATGGGAAGTC 708


RESULT 7
BI176426
LOCUS       BI176426                727 bp    mRNA    linear   EST 07-MAR-2003
DEFINITION  EST517371 cSTE Solanum tuberosum cDNA clone cSTE1A17 5' sequence,
            mRNA sequence.
ACCESSION   BI176426
VERSION     BI176426.1  GI:14642237
KEYWORDS    EST.
```

```
SOURCE      Solanum tuberosum (potato)
  ORGANISM  Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum.
REFERENCE   1  (bases 1 to 727)
  AUTHORS   van der Hoeven,R., Bezzerides,J., Bachem,C., Visser,R., Cho,J.,
            Chiemingo,A., Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and
            Baker,B.
  TITLE     Generation of ESTs from in vitro grown microtubers
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Robin Buell
            The Institute for Genomic Research
            9712 Medical Center Dr, Rockville, MD 20850, USA
            Email: potato-array@tigr.org
            This clone can be obtained from the University of Arizona Genomics
            Institute. Orders can be made through URL:
            http://genome.arizona.edu/orders/
            Seq primer: M13F-R.
FEATURES             Location/Qualifiers
     source          1. .727
                     /organism="Solanum tuberosum"
                     /mol_type="mRNA"
                     /cultivar="Bintje"
                     /db_xref="taxon:4113"
                     /clone="cSTE1A17"
                     /tissue_type="axillary buds of stem explants; growing
                     sink-tubers"
                     /dev_stage="7, 8 and 10 days"
                     /lab_host="SOLR"
                     /clone_lib="cSTE"
                     /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
                     XhoI; Tissue supplied by Christian Bachem and Richard
                     Visser (Department of Plant Breeding, Wageningen
                     University, The Netherlands). The cSTA libraries will
                     attempt to capture the induction and initiation/initial
                     growth of the tuber in an in vitro system as described in
                     Bachem et al. (Plant Journal, 1996). Small microtubers
                     develop from axillary buds attached to stem explants when
                     placed on a high sucrose medium (10%). Visible
                     morphological changes occur synchronously at day five in
                     the axillary buds. The first library, cSTA (1-20) consists
                     of axillary buds harvested on days 1-3. This targets
                     those genes involved in induction of the microtubers. The
                     following libraries, cSTA (21-40) and cSTA (41-60),
                     capture genes involved in tuber initiation and outgrowth.
                     This library is noted as P3 in Tanksley lab notebooks."
ORIGIN

  Query Match             21.7%;  Score 594.8;  DB 4;  Length 727;
  Best Local Similarity   94.9%;  Pred. No. 3.3e-129;
  Matches  695;  Conservative    0;  Mismatches   17;  Indels   20;  Gaps    7;

Qy         17 TATAGATCAGTCTGACAAGAAGGCAACTTCTCAAAGCTTAGAGAGCTACCACCCGAAGAT 76
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db          1 TATAGATCAGTCTGACAAGAAGGCAACTTCTCAAAGTTTAGAGAGCTACCACCCGAAGAT 60

Qy         77 AGACAGTTAGTTACA----------TGTACTGTTATAGATAAAAGGAGAAATCCGAAGAA 126
              |||||||||||||||          ||||||||||||||||||||||||||||||||||
Db         61 AGACAGTTAGTTACATGTGTGTATGTGTACTGTTATAGATAAAAGGAGAAATCCGAAGAA 120

Qy        127 GAAAGAATTTTTTTTGCAGATATGTACTATCAAGGAACCTCGGATAATACTAATATACAA 186
              |||||||||||||||||||||||||||||||||||||||||||||||   ||||||||||
Db        121 GAAAGAATTTTTTTTGCAGATATGTACTATCAAGGAACCTCGGATAA---TAATATACAA 177

Qy        187 GCTGATCATCAACAACGTCATAATCATGGGAATAGTAATAATAATAATATTCAGACACTT 246
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        178 GCTGATCATCAACAACATCATAATCATGGGAATAGTAATAATAATAATATTCAGACACTT 237

Qy        247 TATTTGATGAACCCTAACAATTATATGCAAGGCTACACTACTTCTGACACA---CAGCAG 303
              |||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        238 TATTTGATGAACCCTAACAATTATATGCAAGGCTACACTACTTCTGACACACAGCAGCAG 297

Qy        304 CAGCAGCAGTTACTTTTCCTGAATTCTTCACCAGCAGCAAGCAACGCGCTTTGCCATGCG 363
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        298 CAGCAGCAGTTACTTTTCCTGAATTCTTCACCAGCAGGAAGCAACGCGCTTTGCCATGCG 357

Qy        364 AATATACAACACGCGCCGCTGCAACAGCAGCACTTTGTCGGTGTGCCTCTTCCGGCAGTA 423
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        358 AATATACAACACGCGCCGCTGCAACAGCAGCACTTTGTCGGTGTGCCTCTTCCGGCAGTA 417

Qy        424 AGTTTGCACGATCAGATCAATCATCATGGACTTTTACAGCGCATGTGGAACAACCAAGAT 483
              |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        418 AGTTTGCACGATCAGATCAATCATCATGGACTTTTACAGCGTATGTGGAACAACCAAGAT 477

Qy        484 CAATCTCAGCAGGTGATAGTACCATCGTCGACGGGGGTTTCTGCCACGTCATGTGGCGGG 543
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        478 CAATCTCAGCAGGTGATAGTACCATCGTCGACGGGGGTTTCTGCCACGTCATGTGGCGGG 537

Qy        544 ATCACCACGGAC-TTGGCGTCTCAATTGGCGTTTCAGAGG-CCGATTCCGACACCACAAC 601
              || |||||||| |||||||||||||||||||||||||||| |||||  |   |  |||
Db        538 ATTACCACGGACTTTGGCGTCTCAATTGGCGTTTCAGAGGCCCGATTGTGGTGTCGGAAC 597

Qy        602 ACCGACAGCAGCAACAACAGCAAGGCGGTCTATCTCTAAGCCTTTCTCCTCAGCTACAAC 661
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| || ||
Db        598 ACCGACAGCAGCAACAACAGCAAGGCGGTCTATCTCTAAGCCTTTCTCCTCAGCAAC-AC 656

Qy        662 AGCAAATTAGTTTCAATAACAATATTTCATCCTCATCACCAAGGACAAATAATGTTACTA 721
              |||||||||||||| ||||||||||||||| |||||||||||||||||||||||||||||
Db        657 AGCAAATTAGTTT-AATAACAATATTTCAGCCTCATCACCAAGGACAAATAATGTTACTA 715

Qy        722 TTAGGGGAACAT 733
              |||| |||| ||
Db        716 TTAGAGGAAGAT 727


RESULT 8
AW930221
LOCUS       AW930221                644 bp    mRNA    linear   EST 18-MAY-2001
DEFINITION  EST340678 tomato fruit mature green, TAMU Lycopersicon esculentum
            cDNA clone cLEF41L1 5', mRNA sequence.
ACCESSION   AW930221
VERSION     AW930221.1  GI:8105622
KEYWORDS    EST.
SOURCE      Lycopersicon esculentum (tomato)
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1  (bases 1 to 644)
  AUTHORS   Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F.,
            Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S.,
            Ronning,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and
            Giovannoni,J.
  TITLE     Generation of ESTs from tomato fruit tissue
  JOURNAL   Unpublished (1999)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            5 prime sequence.
FEATURES             Location/Qualifiers
     source          1. .644
                     /organism="Lycopersicon esculentum"
                     /mol_type="mRNA"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cLEF41L1"
                     /tissue_type="fruit pericarp"
                     /dev_stage="mature green (3-5 days pre-ripening)"
                     /lab_host="SOLR"
                     /clone_lib="tomato fruit mature green, TAMU"
                     /note="Vector: pBlueScript SK(-); Site_1: EcoR1; Site_2:
                     Xho1; cLEF - Fruit were tagged at the 1cm stage and
                     harvested 3-5 days prior to  ripening. Fruit were cut in
```

half to verify the seeds were indeed 'immature' and the seeds and locules were discarded prior to freezing the pericarp"

ORIGIN

Query Match 21.3%; Score 581.6; DB 2; Length 644;
Best Local Similarity 95.7%; Pred. No. 4.2e-126;
Matches 622; Conservative 0; Mismatches 19; Indels 9; Gaps 2;

```
Qy       1229 GAGGGAAAATCGAAGGCTCAAGACTCAAATTTGTGGACCATCATCTAAGGCAACAACGCG 1288
             |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAGGGAAAATTGAAGGCTCAAGACTCAAATTTGTGGACCATCATCTAAGGCAACAACGCG 60

Qy       1289 CGCTGCAACAGATAGGAATGATGCAACCAAATGCTTGGAGACCCCAAAGAGGTTTACCTG 1348
             |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CGCTGCAACAGCTAGGAATGATGCAACCAAATGCTTGGAGACCCCAAAGAGGTTTACCTG 120

Qy       1349 AAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTTTTCGAGCATTTTCTTCATCCTTACCCAA 1408
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        121 AAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTTTTCGAGCATTTTCTTCATCCTTACCCCA 180

Qy       1409 AGGATTCAGACAAAATCATGCTTGCTAAGCAAACGGGGCTAACAAGGAGCCAGGTGTCTA 1468
             | || ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db        181 AAGACTCAGACAAAATCATGCTTGCTAAGCAAACGGGGCTAACAAGGAGCCAGGTCTCTA 240

Qy       1469 ACTGGTTCATAAATGCTCGAGTTCGATTATGGAAGCCAATGGTAGAAGAAATGTACTTGG 1528
             |||||||||| ||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        241 ACTGGTTCATTAATGCTCGAGTTCGATTATGGAAGCCAATGGTTGAAGAAATGTACTTGG 300

Qy       1529 AAGAAGTGAAGAATCAAGAACAAAACAGTACTAATACTTCAGGAGATAACAAAAACAAAG 1588
             |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db        301 AAGAAGTGAAGAATCAAGAACAAAACAGTAGTAATACTTCAGGAGATAACAAAAACAAAG 360

Qy       1589 AGACCAATATAAGTGCTCCAAATGAAGAGAAACATCCAATTATTACTAGCAGCTTATTAC 1648
             |||| |||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db        361 AGACGAATATAAGTGCTCCAAATGAAGAGAAACAACCAATTATTACTAGCAGCTTATTAC 420

Qy       1649 AAGATGGTATTACTACTACTCAAGCAGAAATTTCTACCTCAACTATTTCAACTTCCCCTA 1708
             |||| ||      ||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAGAAGG------TACTACTCAAGCAGAAATTTCTACCTCAACTATTTCAACTTCCCCTA 474

Qy       1709 CTGCAGGTGCTTCACTTCATCATGCTCACAATTTCTCCTTCCTTGGTTCATTCAACATGG 1768
             |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        475 CTGCAGGTGCTTCACTTCATCATGCTCACAACTTCTCCTTCCTTGGTTCATTCAACATGG 534

Qy       1769 ATAATACTACTACTACTGTTGATCATATTGAAAACAACGCGAAAAAGCAAAGAA---ATG 1825
             | ||||||||||||||||||||||||||||||||||||||||||||||||| |||||   |||
Db        535 AGAATACTACTACTACTGTTGATCATATTGAAAACAACGCGAAAAAGCCAAGAAATCATG 594

Qy       1826 ACATGCACAAGTTTTCTCCAAGTAGTATTCTTTCATCTGTTGACATGGAA 1875
             |||||||||||||||| |||||| |||||||||||||| ||| ||||||
Db        595 ACATGCACAAGTTTTCTTCAAGTAATATTCTTTCATCTGGTGAAATGGAA 644
```

RESULT 9
BG129304
LOCUS BG129304 630 bp mRNA linear EST 31-JAN-2001
DEFINITION EST474950 tomato shoot/meristem Lycopersicon esculentum cDNA clone cTOF23N12 5' sequence, mRNA sequence.
ACCESSION BG129304
VERSION BG129304.1 GI:12629492
KEYWORDS EST.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 630)
AUTHORS van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Utterback,T., Hansen,C., Ronning,C. and Tanksley,S.
TITLE Generation of ESTs from tomato shoot/meristem tissue
JOURNAL Unpublished (2001)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html.
FEATURES Location/Qualifiers
source 1. .630
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cTOF23N12"
/tissue_type="shoot/meristem"
/dev_stage="developing shoots from 4-6wks old plants"
/lab_host="SOLR"
/clone_lib="tomato shoot/meristem"
/note="Vector: pBluescript SK(-); Site_1: EcoR1; Site_2: Xho1; Small expanding leaves from the growing tip were taken from greenhouse plants (4-6wks old TA496). Tissue was immediately frozen in liquid nitrogen."
ORIGIN

Query Match 20.7%; Score 565; DB 4; Length 630;
Best Local Similarity 95.4%; Pred. No. 3.4e-122;
Matches 606; Conservative 0; Mismatches 20; Indels 9; Gaps 2;

```
Qy       1497 ATGGAAGCCAATGGTAGAAGAAATGTACTTGGAAGAAGTGAAGAATCAAGAACAAAACAG 1556
             |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db          2 ATGGAAGCCAATGGTTGAAGAAATGTACTTGGAAGAAGTGAAGAATCAAGAACAAAACAG 61

Qy       1557 TACTAATACTTCAGGAGATAACAAAAACAAAGAGACCAATATAAGTGCTCCAAATGAAGA 1616
             || |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db         62 TAGTAATACTTCAGGAGATAACAAAAACAAAGAGACGAATATAAGTGCTCCAAATGAAGA 121

Qy       1617 GAAACATCCAATTATTACTAGCAGCTTATTACAAGATGGTATTACTACTACTCAAGCAGA 1676
             ||||||||||||||||||||||||||||||||||      ||||||||||||||||
Db        122 GAAACAACCAATTATTACTAGCAGCTTATTACAAGATGG------TACTACTCAAGCAGA 175

Qy       1677 AATTTCTACCTCAACTATTTCAACTTCCCCTACTGCAGGTGCTTCACTTCATCATGCTCA 1736
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        176 AATTTCTACCTCAACTATTTCAACTTCCCCTACTGCAGGTGCTTCACTTCATCATGCTCA 235

Qy       1737 CAATTTCTCCTTCCTTGGTTCATTCAACATGGATAATACTACTACTACTGTTGATCATAT 1796
             ||| |||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db        236 CAACTTCTCCTTCCTTGGTTCATTCAACATGGAGAATACTACTACTACTGTTGATCATAT 295

Qy       1797 TGAAAACAACGCGAAAAAGCAAAGAA---ATGACATGCACAAGTTTTCTCCAAGTAGTAT 1853
             |||||||||||||||||| |||||   ||||||||||||||||||||||||||||||||
Db        296 TGAAAACAACGCGAAAAAGCCAAGAAATCATGACATGCACAAGTTTTCTCCAAGTAGTAT 355

Qy       1854 TCTTTCATCTGTTGACATGGAAGCCAAAGCTAGAGAATCATCAAATAAAGGGTTTACTAA 1913
             |||||||||||||||| |||||||||||||||||||||||| |||| ||||||||||||||
Db        356 TCTTTCATCTGTTGAAATGGAAGCCAAAGCTAGAGAATCAACAAACAAAGGGTTTACTAA 415

Qy       1914 TCCTTTAATGGCAGCATACGCGATGGGAGATTTTGGAAGGTTTGATCCTCATGATCAACA 1973
             |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db        416 TCCTTTAATGGCAGCGTACGCGATGGGAGATTTTGGAAGGTTTGATCCTCATGATCAACA 475

Qy       1974 AATGACCGCGAATTTTCATGGAAATAATGGTGTCTCTCTTACTTTAGGACTTCCTCCTTC 2033
             |||||||||||||||||||||||||||||||| |||||||||||||||||||| ||||
Db        476 AATGACCGCGAATTTTCATGGAAATAATGGTGTATCTCTTACTTTAGGACTTCCTACTTC 535

Qy       2034 TGAAAACCTAGCCATGCCAGTGAGCCAACAAAATTACCTTTCTAATGACTTGGGAAGTAG 2093
             |||||||||||||||||||||||||||||||||||||||| |||||||||| ||| ||||||
Db        536 TGAAAACCTAGCCATGCCAGTGAGCCAACAAAATTATCTTTCTAATGAGTTGCGAAGTAT 595

Qy       2094 GTCTGAAATGGGGAGTCATTACAATAGAATGGGAT 2128
             | ||||||| || ||||||||||||||||||||||
Db        596 GCCTGAAATAGGAAGTCATTACAATAGAATGGGAT 630
```

RESULT 10
BI921008
LOCUS BI921008 789 bp mRNA linear EST 10-MAR-2003

```
DEFINITION  EST540943 potato microtubers, in vitro-grown Solanum tuberosum cDNA
            clone cSTE27I21 5' end, mRNA sequence.
ACCESSION   BI921008
VERSION     BI921008.1  GI:16217036
KEYWORDS    EST.
SOURCE      Solanum tuberosum (potato)
  ORGANISM  Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum.
REFERENCE   1  (bases 1 to 789)
  AUTHORS   van der Hoeven,R., Bezzerides,J., Bachem,C., Visser,R.,
            Karamycheva,S.A., Tsai,J., Van Aken,S., Utterback,T., Chiemingo,A.,
            Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.
  TITLE     Generation of ESTs from in vitro grown microtubers (2001b)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Robin Buell
            The Institute for Genomic Research
            9712 Medical Center Dr, Rockville, MD 20850, USA
            Email: potato-array@tigr.org
            This clone can be obtained from the University of Arizona Genomics
            Institute. Orders can be made through URL:
            http://genome.arizona.edu/orders/
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1. .789
                     /organism="Solanum tuberosum"
                     /mol_type="mRNA"
                     /cultivar="Bintje"
                     /db_xref="taxon:4113"
                     /clone="cSTE27I21"
                     /tissue_type="axillary buds of stem explants; growing
                     sink-tubers"
                     /dev_stage="7, 8 and 10 days"
                     /lab_host="SOLR"
                     /clone_lib="potato microtubers, in vitro-grown"
                     /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
                     XhoI; Tissue supplied by Christian Bachem and Richard
                     Visser (Department of Plant Breeding, Wageningen
                     University, The Netherlands); sequencing by The Institute
                     for Genomic Research. The cSTA libraries will attempt to
                     capture the induction and initiation/initial growth of the
                     tuber in an in vitro system as described in Bachem et al.
                     (Plant Journal, 1996). Small microtubers develop from
                     axillary buds attached to stem explants when placed on a
                     high sucrose medium (10%). Visible morphological changes
                     occur synchronously at day five in the axillary buds. The
                     first library, cSTA (1-20) consists of axillary buds
                     harvested on days 1-3. This targets those genes involved
                     in induction of the microtubers. The following libraries,
                     cSTA (21-40) and cSTA (41-60), capture genes involved in
                     tuber initiation and outgrowth. This library is noted as
                     P3  in Tanksley lab notebooks."
ORIGIN

  Query Match             20.6%;  Score 564.2;  DB 4;  Length 789;
  Best Local Similarity   99.5%;  Pred. No. 5.6e-122;
  Matches  566;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          87 TTACATGTACTGTTATAGATAAAAGGAGAAATCCGAAGAAGAAAGAATTTTTTTTGCAGA 146
               || || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         221 TTGCAGGTACTGTTATAGATAAAAGGAGAAATCCGAAGAAGAAAGAATTTTTTTTGCAGA 280

Qy         147 TATGTACTATCAAGGAACCTCGGATAATACTAATATACAAGCTGATCATCAACAACGTCA 206
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db         281 TATGTACTATCAAGGAACCTCGGATAATACTAATATACAAGCTGATCATCAACAACATCA 340

Qy         207 TAATCATGGGAATAGTAATAATAATAATATTCAGACACTTTATTTGATGAACCCTAACAA 266
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         341 TAATCATGGGAATAGTAATAATAATAATATTCAGACACTTTATTTGATGAACCCTAACAA 400

Qy         267 TTATATGCAAGGCTACACTACTTCTGACACACAGCAGCAGCAGCAGTTACTTTTCCTGAA 326
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         401 TTATATGCAAGGCTACACTACTTCTGACACACAGCAGCAGCAGCAGTTACTTTTCCTGAA 460

Qy         327 TTCTTCACCAGCAGCAAGCAACGCGCTTTGCCATGCGAATATACAACACGCGCCGCTGCA 386
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         461 TTCTTCACCAGCAGCAAGCAACGCGCTTTGCCATGCGAATATACAACACGCGCCGCTGCA 520

Qy         387 ACAGCAGCACTTTGTCGGTGTGCCTCTTCCGGCAGTAAGTTTGCACGATCAGATCAATCA 446
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         521 ACAGCAGCACTTTGTCGGTGTGCCTCTTCCGGCAGTAAGTTTGCACGATCAGATCAATCA 580

Qy         447 TCATGGACTTTTACAGCGCATGTGGAACAACCAAGATCAATCTCAGCAGGTGATAGTACC 506
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         581 TCATGGACTTTTACAGCGCATGTGGAACAACCAAGATCAATCTCAGCAGGTGATAGTACC 640

Qy         507 ATCGTCGACGGGGGTTTCTGCCACGTCATGTGGCGGGATCACCACGGACTTGGCGTCTCA 566
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         641 ATCGTCGACGGGGGTTTCTGCCACGTCATGTGGCGGGATCACCACGGACTTGGCGTCTCA 700

Qy         567 ATTGGCGTTTCAGAGGCCGATTCCGACACCACAACACCGACAGCAGCAACAACAGCAAGG 626
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         701 ATTGGCGTTTCAGAGGCCGATTCCGACACCACAACACCGACAGCAGCAACAACAGCAAGG 760

Qy         627 CGGTCTATCTCTAAGCCTTTCTCCTCAGC 655
               |||||||||||||||||||||||||||||
Db         761 CGGTCTATCTCTAAGCCTTTCTCCTCAGC 789


RESULT 11
CO909070
LOCUS       CO909070                610 bp    mRNA    linear   EST 16-AUG-2004
DEFINITION  BJ02039F01 BJ02 Capsicum annuum cDNA 5', mRNA sequence.
ACCESSION   CO909070
VERSION     CO909070.1  GI:51299373
KEYWORDS    EST.
SOURCE      Capsicum annuum
  ORGANISM  Capsicum annuum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Capsicum.
REFERENCE   1  (bases 1 to 610)
  AUTHORS   Song,W.Y., Jeon,W.B., Kim,K.S., Lee,H.H., Ko,M.K., Kim,Y.S.,
            Hong,J.C. and Oh,B.J.
  TITLE     Generation of Hot Pepper (Capsicum annuum) ESTs (Express Sequence
            Tags) from Red Ripe Fruit (Song, et al.)
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Boung-Jun Oh
            Jeonnam Biotechnology Research Center
            Namyang Bld. #603, 10-4 Gwangsan-dong, Dong-gu, Gwangju 501-180,
            Korea
            Tel: 82 62 607 2422
            Fax: 82 62 607 6205
            Email: bjo@biohub.re.kr
            Plate: 039  row: F  column: 01.
FEATURES             Location/Qualifiers
     source          1. .610
                     /organism="Capsicum annuum"
                     /mol_type="mRNA"
                     /cultivar="Nokkwang"
                     /db_xref="taxon:4072"
                     /tissue_type="red pepper fruit pericarp"
                     /clone_lib="BJ02"
                     /note="Vector: pBluescript II SK(+)XR; Site_1: EcoRI;
                     Site_2: XhoI; cDNA library was generated from red ripe
                     fruit pericarp using lambda Zap II phage vector. In vivo
                     excision was done with helper phage to generate subclone
                     in pBluescript II SK(+)XR] vector."
ORIGIN

  Query Match             19.4%;  Score 531;  DB 7;  Length 610;
  Best Local Similarity   92.5%;  Pred. No. 3.6e-114;
  Matches  558;  Conservative    0;  Mismatches   45;  Indels    0;  Gaps    0;

Qy         911 CTGGTGGTGGTGAAAGTAGCAGCAGGCAGAAAAATGAAGTTGCTGTTGAGCTTACAACTG 970
```

```
                || || |||||    |||||||||||| |||||||||||||||| ||||||||||||||
Db          8 CTAGTAGTGGTACTGGTAGCAGCAGGCCGAAAAATGAAGTTGCTGCTGAGCTTACAACTG 67

Qy        971 CTCAAAGACAAGAACTTCAAATGAAAAAAGCCAAGCTTCTTGCCATGCTTGAAGAGGTGG 1030
              |||||||||||||||||| |||||||| || |||||| |  ||||||||||||||||||
Db         68 CTCAAAGACAAGAACTTCAGATGAAAAAGGCAAAGCTTGTCACCATGCTTGAAGAGGTGG 127

Qy       1031 AGCAAAGGTACAGACAGTACCATCACCAAATGCAAATAATTGTATTATCATTTGAGCAAG 1090
              | ||||||||||||||||||||||||||||||||||||||||||| | ||||||||||||||
Db        128 AACAAAGGTACAGACAGTACCATCACCAAATGCAAATAATTGTGTCATCATTTGAGCAAG 187

Qy       1091 TAGCAGGAATTGGATCAGCCAAATCATACACTCAATTAGCTTTGCATGCAATTTCGAAGC 1150
              |||||||||||||||||||||||||||||||||||| | ||||||||||||||||||||
Db        188 TAGCAGGAATTGGATCAGCCAAATCATACACTCAACTGGCTTTGCATGCAATTTCGAAGC 247

Qy       1151 AATTCAGATGCCTAAAGGATGCAATTGCTGAGCAAGTAAAGGCGACGAGCAAGAGTTTAG 1210
              ||||| |||| ||||| |||||||||||||| ||| | ||||||||||||||||||||||
Db        248 AATTCCGATGTCTAAAAGATGCAATTGCTGAACAAATTAAGGCGACGAGCAAGAGTTTAG 307

Qy       1211 GTGAAGAGGAAGGCTTGGGAGGGAAAATCGAAGGCTCAAGACTCAAATTTGTGGACCATC 1270
              |||||||||||||| | |||||||||||| ||||| || |||||||||||||| |||||||
Db        308 GTGAAGAGGAAGGCATAGGAGGGAAAATTGAAGGGTCGAGACTCAAATTTGTAGACCATC 367

Qy       1271 ATCTAAGGCAACAACGCGCGCTGCAACAGATAGGAATGATGCAACCAAATGCTTGGAGAC 1330
              |||||||||||||||| ||||||||||  |||||||||||||||||||||||||||||||
Db        368 ATCTAAGGCAACAACGTGCGCTGCAACAATTAGGAATGATGCAACCAAATGCTTGGAGAC 427

Qy       1331 CCCAAAGAGGTTTACCTGAAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTTTTCGAGCATT 1390
              | |||||||||  |||||||||  |||||||| |||||||||||||||||||||||||||
Db        428 CTCAAAGAGGTCTACCTGAAAGGGCTGTCTCCGTCCTTCGTGCTTGGCTTTTCGAGCATT 487

Qy       1391 TTCTTCATCCTTACCCAAAGGATTCAGACAAAATCATGCTTGCTAAGCAAACGGGGCTAA 1450
              |||||||||||||||| |||||||||||||||||||||||||||||||||| || ||||
Db        488 TTCTTCATCCTTACCCCAAGGATTCAGACAAAATCATGCTTGCTAAGCAAACCGGACTAA 547

Qy       1451 CAAGGAGCCAGGTGTCTAACTGGTTCATAAATGCTCGAGTTCGATTATGGAAGCCAATGG 1510
              |||||||||||||| || |||||||||||||||||||||||||||| ||||||||||||||
Db        548 CAAGGAGCCAGGTCTCCAACTGGTTCATAAATGCTCGAGTTCGACTATGGAAGCCAATGG 607

Qy       1511 TAG 1513
              | ||
Db        608 TGG 610
```

RESULT 12
AI773253

```
LOCUS       AI773253                553 bp    mRNA    linear   EST 18-MAY-2001
DEFINITION  EST254353 tomato resistant, Cornell Lycopersicon esculentum cDNA
            clone cLER5L8, mRNA sequence.
ACCESSION   AI773253
VERSION     AI773253.1  GI:5271294
KEYWORDS    EST.
SOURCE      Lycopersicon esculentum (tomato)
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1  (bases 1 to 553)
  AUTHORS   D' Ascenzo,M., He,X., Lyman,J., Matern,A.L., Vision,T., Holt,I.E.,
            Liang,F., Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.Y.,
            Bowman,C.L., Nierman,W., Fraser,C.M., Venter,J.C., Tanksley,S.D.,
            Giovannoni,J.J. and Martin,G.B.
  TITLE     Generation of ESTs from Pseudomonas resistant tomato
  JOURNAL   Unpublished (1999)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            5 prime sequence.
FEATURES             Location/Qualifiers
     source          1. .553
                     /organism="Lycopersicon esculentum"
                     /mol_type="mRNA"
                     /cultivar="R11-12 (35S::Pto in Rio Grande x Money Maker)"
                     /db_xref="taxon:4081"
                     /clone="cLER5L8"
                     /tissue_type="leaf"
                     /dev_stage="4-week old"
                     /lab_host="SOLR"
                     /clone_lib="tomato resistant, Cornell"
                     /note="Vector: pBlueScript SK(-); Site_1: EcoR1; Site_2:
                     Xho1; cLER - Tomato Pseudomonas Resistant EST Library.
                     Directionally cloned  cDNAs inserted into pBlueScript
                     SK(-) at 5' end with EcoRI and 3' end  with XhoI site."
ORIGIN
```

```
  Query Match            18.9%;  Score 517.8;  DB 1;  Length 553;
  Best Local Similarity  96.0%;  Pred. No. 4.6e-111;
  Matches  531;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy       1094 CAGGAATTGGATCAGCCAAATCATACACTCAATTAGCTTTGCATGCAATTTCGAAGCAAT 1153
              ||||| |||||||||| ||||||||||||||||||||||||||||||||| |||||||
Db          1 CAGGAGTTGGATCAGCAAAATCATACACTCAATTAGCTTTGCATGCAATTTCAAAGCAAT 60

Qy       1154 TCAGATGCCTAAAGGATGCAATTGCTGAGCAAGTAAAGGCGACGAGCAAGAGTTTAGGTG 1213
              |||||||||||||||||||||  ||||||||| ||||||||||||||||||||||||||
Db         61 TCAGATGCCTAAAGGATGCAATTAGTGAGCAAGTGAAGGCGACGAGCAAGAGTTTAGGTG 120

Qy       1214 AAGAGGAAGGCTTGGGAGGGAAAATCGAAGGCTCAAGACTCAAATTTGTGGACCATCATC 1273
              |||| |||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        121 AAGATGAAGGCTTGGGAGGGAAAATTGAAGGCTCAAGACTCAAATTTGTGGACCATCATC 180

Qy       1274 TAAGGCAACAACGCGCGCTGCAACAGATAGGAATGATGCAACCAAATGCTTGGAGACCCC 1333
              |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db        181 TAAGGCAACAACGCGCGCTGCAACAGCTAGGAATGATGCAACCAAATGCTTGGAGACCCC 240

Qy       1334 AAAGAGGTTTACCTGAAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTTTTCGAGCATTTTC 1393
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAAGAGGTTTACCTGAAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTTTTCGAGCATTTTC 300

Qy       1394 TTCATCCTTACCCAAAGGATTCAGACAAAATCATGCTTGCTAAGCAAACGGGGCTAACAA 1453
              |||||||||||| || |||||||||||||||||||||||||||||||||||||||||||
Db        301 TTCATCCTTACCCCAAAGACTCAGACAAAATCATGCTTGCTAAGCAAACGGGGCTAACAA 360

Qy       1454 GGAGCCAGGTGTCTAACTGGTTCATAAATGCTCGAGTTCGATTATGGAAGCCAATGGTAG 1513
              |||||||||| |||||||||||||| ||||||||||||||||||||||||||||||| |
Db        361 GGAGCCAGGTCTCTAACTGGTTCATTAATGCTCGAGTTCGATTATGGAAGCCAATGGTTG 420

Qy       1514 AAGAAATGTACTTGGAAGAAGTGAAGAATCAAGAACAAAACAGTACTAATACTTCAGGAG 1573
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        421 AAGAAATGTACTTGGAAGAAGTGAAGAATCAAGAACAAAACAGTAGTAATACTTCAGGAG 480

Qy       1574 ATAACAAAAACAAAGAGACCAATATAAGTGCTCCAAATGAAGAGAAACATCCAATTATTA 1633
              |||||||||||||||||| ||  | |||||||||||||||||||| || |||||||||||
Db        481 ATAACAAAAACAAAGAGACGAAATTTAGTGCTCCAAATGAAGAGAAACAACCCATTATTA 540

Qy       1634 CTAGCAGCTTATT 1646
              |||||||||||||
Db        541 CTAGCAGCTTATT 553
```

RESULT 13
BE435130/c

```
LOCUS       BE435130                721 bp    mRNA    linear   EST 18-MAY-2001
DEFINITION  EST406208 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
            clone cLEG25M5, mRNA sequence.
ACCESSION   BE435130
VERSION     BE435130.1  GI:9432973
KEYWORDS    EST.
SOURCE      Lycopersicon esculentum (tomato)
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1  (bases 1 to 721)
```

```
  AUTHORS   Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,
            Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M.,
            Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and
            Tanksley,S.D.
  TITLE     Generation of ESTs from tomato fruit tissue, breaker stage
  JOURNAL   Unpublished (2000)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            5 prime sequence.
FEATURES             Location/Qualifiers
     source          1. .721
                     /organism="Lycopersicon esculentum"
                     /mol_type="mRNA"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cLEG25M5"
                     /tissue_type="Pericarp"
                     /dev_stage="breaker"
                     /lab_host="SOLR"
                     /clone_lib="tomato breaker fruit, TIGR"
                     /note="Vector: pBluescriptSKmCUadapt; Site_1: EcoR1;
                     Site_2: XhoI; Fruit were harvested at the breaker stage
                     (first sign of lycopene accumulation on the blossom end of
                     the fruit).  Fruit were cut in half and the seeds and
                     locules were discarded prior to freezing the pericarp."
ORIGIN

  Query Match             18.4%;  Score 502.4;  DB 2;  Length 721;
  Best Local Similarity   88.6%;  Pred. No. 2.1e-107;
  Matches  654;  Conservative    0;  Mismatches   31;  Indels   53;  Gaps    8;

Qy       2013 TACTTTAGGACTTCCTCCTTCTGAAAACCTAGCCATGCCAGTGAGCCAACAAAATTACCT 2072
             ||||||||||||| ||| ||||||||||||||||||||||||||||||||||| || ||
Db        721 TACTTTAGGACTT-CTCTTTCTGAAAACCTAGCCATGCCAGTGAGCCAACAAAA-TATCT 664

Qy       2073 TTCTAATGACTTGGGAAGTAGGTCTGAAATGGGGAGTCATTACAATAGAATGGGATATGA 2132
             ||||||||| |||||||||||| ||||||| || ||||||||||||||||||||||||||
Db        663 TTCTAATGAGTTGGGAAGTAGGCCTGAAATAGGAAGTCATTACAATAGAATGGGATATGA 604

Qy       2133 AAACATTGATTTTCAGAGTGGGAATAAGCGATTTCCGACTCAACTATTACCAGATTTTGT 2192
             |||||||||||||| |||||||||||||||||| |||||||||||| ||||||||||||
Db        603 AAACATTGATTTTCAAAGTGGGAATAAGCGATTTCCAACTCAACTATTGCCAGATTTTGT 544

Qy       2193 TACAGGTAATCTAGGAACATGAATACCAGAAAGTCTCGTATTGATAGCTGAAAAGAT--A 2250
             |||||||||||| |||||||||||||| |||||||||||||||||||||||||| ||  |
Db        543 TACAGGTAATCTTGGAACATGAATACCACAAAGTCTCGTATTGATAGCTGAAAACATAAA 484

Qy       2251 AAAGGAAGTTAGGGATACTCTTATATTGTGTGAGGCCTTCTGGCCCAAGTCGGAGGACCC 2310
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 AAAGGAAGTTAGGGATACTCTTATATTGTGTGAGGCCTTCTGGCCCAAGTCGGAGGACCC 424

Qy       2311 AATTTGATACAACCTATCATAGGAGAAAAGAAGTGGAGACTAAATTAAAGTAACAAAATT 2370
             |||||||||||||||||||||||||                              |||||| |
Db        423 AATTTGATACAACCTATCATAGGAG--------------------------ACAAAAGT 391

Qy       2371 TTAAAGCACACTTTCTAGTATATATACTTCTTTTTT---TTATAGTATAGAAAAGAAGAG 2427
             |||||||||||||||||||||||||||||||||||   ||||||||||||||||||||||
Db        390 TTAAAGCACACTTTCTAGTATATATACTTCTTTTTTTAATTATAGTATAGAAAAGAAGAG 331

Qy       2428 ATTTTGTGCTTTAGTGTATAGATAGAGTCTACTTAGTATAGGTTATACTTCTAGTTCCTT 2487
             |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        330 ATTTTGTGCTTTAGTGTATAGATAGAGTCTACTTAGTATAGGTTATACTTCTACTTCCTT 271

Qy       2488 GAGAAGATTGATACAACTAGTAGTATTTTTTTTCTTTTGG----------GTTGGCTTGG 2537
             |||||||||||||||||||||||||||||||| || ||  ||          | | ||||||
Db        270 GAGAAGATTGATACAACTAGTAGTATTTTTATTTTTGGGGGGTTGAAATAGCTTGCTTGG 211

Qy       2538 AGTACTATTTTAAGTTAT-------TGGAAACTAGCTATAG--TAAATGTTGTAAAGTTG 2588
             |||||||||| |  | ||       |  ||||||||||||  || ||||||||||||||
Db        210 AGTACTATTTAAGTTAATGTGGAAATTTAAACTAGCTATAGTATATATGTTGTAAAGTTG 151

Qy       2589 TGATATTGTTCCTCTCAATTTGCATATAATTTGAAATATTTTGTACCTACTAGCTAGTCT 2648
             |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        150 TGATATTGTTACTCTCAATTTGCATATAATTTGAAATATTTTGTACCTACTAGCTAGTCT 91

Qy       2649 CTAAATTATGTTTCCATTGCTTGTAATTGCAATTTTATTTGAATTTTGTGCTATCATTAT 2708
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          90 CTAAATTATGTTTCCATTGCTTGTAATTGCAATTTTATTTGAATTTTGTGCTATCATTAT 31

Qy       2709 TAGATTAGCAAAAAAAAA 2726
             ||||||||||||  | |
Db          30 TAGATTAGCAAAAGTACA 13


RESULT 14
CK262102
LOCUS       CK262102                964 bp    mRNA    linear   EST 03-AUG-2004
DEFINITION  EST708180 potato abiotic stress cDNA library Solanum tuberosum cDNA
            clone POABE14 5' end, mRNA sequence.
ACCESSION   CK262102
VERSION     CK262102.1  GI:39819080
KEYWORDS    EST.
SOURCE      Solanum tuberosum (potato)
  ORGANISM  Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum.
REFERENCE   1  (bases 1 to 964)
  AUTHORS   Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.
  TITLE     Generation of ESTs from abiotic stressed potato tissue
  JOURNAL   Unpublished (2003)
COMMENT     Other_ESTs: EST708181
            Contact: Robin Buell
            The Institute for Genomic Research
            9712 Medical Center Dr, Rockville, MD 20850, USA
            Email: potato-array@tigr.org
            Clones can be requested from the University of Arizona Genomics
            Institute via http://genome.arizona.edu/orders/ .
            Seq primer: ATT TAG GTG ACA CTA TAG.
FEATURES             Location/Qualifiers
     source          1. .964
                     /organism="Solanum tuberosum"
                     /mol_type="mRNA"
                     /cultivar="Kennebec"
                     /db_xref="taxon:4113"
                     /clone="POABE14"
                     /tissue_type="abiotic stress treated leaf and root tissue"
                     /lab_host="DH10B-TonA"
                     /clone_lib="potato abiotic stress cDNA library"
                     /note="Vector: pCMVSport6.1; Site_1: EcoRI; Site_2: NotI;
                     supplier: Solanum tuberosum var. Kennebec plants were
                     grown from cuttings on a 16hr light/8 hr dark cycle at 25
                     C for 3-4 weeks. Abiotic stress conditions were applied to
                     four separate sets of plants. Set 1 involved saturation of
                     the soil with 150 mM NaCl and tissues were harvested at
                     following application of the salt stress (leaves: 2hr,
                     6hr, 12hr, 1d, 2d, and 4d; roots:2hr, 6hr, 12hr, and 2d).
                     Set 2 were grown under the standard conditions and then
                     were water stressed by withdrawal of further watering
                     applications. Drought stressed plants were harvested after
                     cessation of watering (leaves: 3d, 5d, and 7d; roots:3d
                     and 5d). Set 3 were grown under the standard conditions
                     and then were cold stressed by placement at 4 C. Cold
                     stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
                     and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
                     2d. Set 4 were grown under the standard conditions and
                     then were heat stressed by placement at 35 C. Heat
                     stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
                     2d and 4d and heat-stressed roots were harvested at 6 hr,
                     12 hr, 1 d, and 4d. RNA was isolated from all tissues and
                     equal RNA from each tissue and stress was pooled to
                     construct the cDNA library. RNA sample."
ORIGIN
```

```
  Query Match           17.4%;  Score 476;  DB 7;  Length 964;
  Best Local Similarity  72.6%;  Pred. No. 3.8e-101;
  Matches  687;  Conservative    0;  Mismatches  220;  Indels   39;  Gaps    4;

Qy        909 TTCTGGTGGTGGTGAAAGTAGCAGCAGGCAGAAAAATGAAGTTGCTGTTGAGCTTACAAC 968
              ||  | | ||| |||    |   || ||      ||||  |  ||| ||||| ||||| ||
Db         34 TTTGGCTAGTGATGATGTTGAAAGTAGCAGCCAAAAAAATATTGTTGTTGAACTTACTAC 93

Qy        969 TGCTCAAAGACAAGAACTTCAAATGAAAAAAGCCAAGCTTCTTGCCATGCTTGAAGAGGT 1028
              ||||||||||||||||||||||||| ||||||||||| ||  ||||||||| |||||
Db         94 AGCTCAAAGACAAGAACTTCAAATGAAGAAAGCCAAGCTTGTTAGCATGCTTGATGAGGT 153

Qy        1029 GGAGCAAAGGTACAGACAGTACCATCACCAAATGCAAATAATTGTATTATCATTTGAGCA 1088
               ||| ||||||||| ||||| |||||||||||||||||||| |||| |  ||||||||||||
Db         154 GGATCAAAGGTATAGACAATACCATCACCAAATGCAAATGATTGCAACATCATTTGAGCA 213

Qy        1089 AGTAGCAGGAATTGGATCAGCCAAATCATACACTCAATTAGCTTTGCATGCAATTTCGAA 1148
               |  | ||||||||||||||| | ||||||||||| ||| | ||||||||  ||||||||||
Db         214 AACAACAGGAATTGGATCATCAAAATCATACACACAACTTGCTTTGCACACAATTTCGAA 273

Qy        1149 GCAATTCAGATGCCTAAAGGATGCAATTGCTGAGCAAGTAAAGGCGACGAGCAAGAGTTT 1208
               |||||| |||||  |||| ||||||||| ||| |||| ||||||  || || || | |||
Db         274 GCAATTTAGATGTTTAAAAGATGCAATTTCTGGGCAAATAAAGGACACAAGTAAAACTTT 333

Qy        1209 AGGTGAAGAGGAAGGCTTGGGAGGGAAAATCGAAGGCTCAAGACTCAAATTTGTGGACCA 1268
               ||| ||||| ||    | | |||||  ||||| ||||| ||||    | ||||||||||| ||
Db         334 AGGGGAAGAAGAGAACATTGGAGGCAAAATTGAAGGATCAAAGTTGAAATTTGTGGATCA 393

Qy        1269 TCATCTAAGGCAACAACGCGCGCTGCAACAGATAGGAATGATGCAACCAAATGCTTGGAG 1328
               |||| || | |||||||| || || |||||  |||| |||||||| | ||||| |||||
Db         394 TCATTTACGCCAACAACGTGCACTACAACAATTAGGGATGATGCAAACCAATGCATGGAG 453

Qy        1329 ACCCCAAAGAGGTTTACCTGAAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTTTTCGAGCA 1388
               || || |||||||| || |||||||| || || || || ||||||||||||||||||||
Db         454 GCCACAGAGAGGTTTGCCCGAAAGAGCGGTTTCGGTTCTCCGCGCTTGGCTTTTCGAGCA 513

Qy        1389 TTTTCTTCATCCTTACCCAAAGGATTCAGACAAAATCATGCTTGCTAAGCAAACGGGGCT 1448
               |||||||||||| || || || |||||||| |||||||||||||||||||||||||| |||||
Db         514 TTTTCTTCATCCGTATCCCAAAGATTCAGATAAAATCATGCTTGCTAAGCAAACAGGGCT 573

Qy        1449 AACAAGGAGCCAGGTGTCTAACTGGTTCATAAATGCTCGAGTTCGATTATGGAAGCCAAT 1508
               |||||||||||||| || || ||||| |||||||| ||||| || ||||||||||||||
Db         574 AACAAGGAGCCAGGTATCAAATTGGTTTATAAATGCTAGAGTTAGACTATGGAAGCCAAT 633

Qy        1509 GGTAGAAGAAATGTACTTGGAAGAAGTGAAGAATCAAGAACAAAACAGTACTAATACTTC 1568
               |||||||||||||||| ||||||||||||||||     | ||| |    | || |    |
Db         634 GGTAGAAGAAATGTACATGGAAGAAGTGAAGAAAACCAATCAAGAACAAAATATTGAGCC 693

Qy        1569 AGGAGATAACAAAAACAAAGAGACCAATATAAGTGCTCCAAATGAAGAGAAACATCCAAT 1628
                    | ||  |||      |   | ||   ||||| ||||    ||||||||   |||||
Db         694 TAATAACAATGAAATTGTTGGTTCAAAATCAAGTGTTCCA---CAAGAGAAATTACCAAT 750

Qy        1629 TATTACTAGCAGCTTATTACAAGATGGTATTACTACTACTCAAGCAGAAATTTCTACCTC 1688
               || ||  |          |   |  |  || | || ||| |   || ||||||||  ||
Db         751 TAGTAGCA-----------ATATTATTCATAATGCTTCTCCAAATGATATTTCTACTTC 798

Qy        1689 AACTATTTCAACTTCCCCTACTGCAGGTGCTTCACTTCATCATGCTCACA---------A 1739
               || |||||||| || || ||  || |           ||  |||||  |
Db         799 CACCATTTCAACATCTCCGACAGGCGGCGGCGGTTCGATTCCGGCTCAGACGGTTGCAGG 858

Qy        1740 TTTCTCCTTCCTTGGTTCATTCAACATGGATAATACTACTACTACTGTTGATCATATTGA 1799
               |||||||||| || | ||||| |||||||| || |                | ||  | |
Db         859 TTTCTCCTTCATTAGGTCATTAAACATGGAGAACA--------------TTGATGATCA 903

Qy        1800 AAACAACGCGAAAAAGCAAAGAAATGACATGCACAAGTTTTCTCCA 1845
               ||  |||   ||||||  ||||||||| ||||| || | ||| | |
Db         904 AAGGAACAACAAAAAGGCAAGAAATGAGATGCAAAATTGTTCACTA 949


RESULT 15
BE460264
LOCUS       BE460264                495 bp    mRNA    linear   EST 18-MAY-2001
DEFINITION  EST411755 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
            clone cLEG28B4, mRNA sequence.
ACCESSION   BE460264
VERSION     BE460264.1  GI:9504638
KEYWORDS    EST.
SOURCE      Lycopersicon esculentum (tomato)
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1  (bases 1 to 495)
  AUTHORS   Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,
            Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M.,
            Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and
            Tanksley,S.D.
  TITLE     Generation of ESTs from tomato fruit tissue, breaker stage
  JOURNAL   Unpublished (2000)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            5 prime sequence.
FEATURES             Location/Qualifiers
     source          1. .495
                     /organism="Lycopersicon esculentum"
                     /mol_type="mRNA"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cLEG28B4"
                     /tissue_type="Pericarp"
                     /dev_stage="breaker"
                     /lab_host="SOLR"
                     /clone_lib="tomato breaker fruit, TIGR"
                     /note="Vector: pBluescriptSKmCUadapt; Site_1: EcoR1;
                     Site_2: XhoI; Fruit were harvested at the breaker stage
                     (first sign of lycopene accumulation on the blossom end of
                     the fruit).  Fruit were cut in half and the seeds and
                     locules were discarded prior to freezing the pericarp."
ORIGIN

  Query Match           17.1%;  Score 467.8;  DB 2;  Length 495;
  Best Local Similarity  96.6%;  Pred. No. 2.9e-99;
  Matches  478;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy        1042 AGACAGTACCATCACCAAATGCAAATAATTGTATTATCATTTGAGCAAGTAGCAGGAATT 1101
               ||||||| |||||||||||||||||||||||| |||||||||||||||||||||||| ||
Db           1 AGACAGTCCCATCACCAAATGCAAATAATTGTATCATCATTTGAGCAAGTAGCAGGAGTT 60

Qy        1102 GGATCAGCCAAATCATACACTCAATTAGCTTTGCATGCAATTTCGAAGCAATTCAGATGC 1161
               |||||||| |||||||||||||||||||||||||||||||||| ||||||||||||||||
Db          61 GGATCAGCAAAATCATACACTCAATTAGCTTTGCATGCAATTTCAAAGCAATTCAGATGC 120

Qy        1162 CTAAAGGATGCAATTGCTGAGCAAGTAAAGGCGACGAGCAAGAGTTTAGGTGAAGAGGAA 1221
               ||||||||||||||  |||||||| |||||||||||||||||||||||||||||| |||
Db          121 CTAAAGGATGCAATTAGTGAGCAAGTGAAGGCGACGAGCAAGAGTTTAGGTGAAGATGAA 180

Qy        1222 GGCTTGGGAGGGAAAATCGAAGGCTCAAGACTCAAATTTGTGGACCATCATCTAAGGCAA 1281
               |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db          181 GGCTTGGGAGGGAAAATTGAAGGCTCAAGACTCAAATTTGTGGACCATCATCTAAGGCAA 240

Qy        1282 CAACGCGCGCTGCAACAGATAGGAATGATGCAACCAAATGCTTGGAGACCCCAAAGAGGT 1341
               |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db          241 CAACGCGCGCTGCAACAGCTAGGAATGATGCAACCAAATGCTTGGAGACCCCAAAGAGGT 300

Qy        1342 TTACCTGAAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTTTTCGAGCATTTTCTTCATCCT 1401
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          301 TTACCTGAAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTTTTCGAGCATTTTCTTCATCCT 360

Qy        1402 TACCCAAAGGATTCAGACAAAATCATGCTTGCTAAGCAAACGGGGCTAACAAGGAGCCAG 1461
               ||||| || || ||||||||||||||||||||||||||||||||||||||||||||||||
Db          361 TACCCCAAAGACTCAGACAAAATCATGCTTGCTAAGCAAACGGGGCTAACAAGGAGCCAG 420
```

```
Qy       1462 GTGTCTAACTGGTTCATAAATGCTCGAGTTCGATTATGGAAGCCAATGGTAGAAGAAATG 1521
              || |||||||||||||| |||||||||||||||||||||||||||||| |||||||||
Db        421 GTCTCTAACTGGTTCATTAATGCTCGAGTTCGATTATGGAAGCCAATGGTTGAAGAAATG 480

Qy       1522 TACTTGGAAGAAGTG 1536
              |||||||||||||||
Db        481 TACTTGGAAGAAGTG 495
```

Search completed: July 29, 2005, 14:13:05
Job time : 8595 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 29, 2005, 11:50:07 ; Search time 8001 Seconds
(without alignments)
4166.628 Million cell updates/sec

Title: US-10-624-201A-2
Perfect score: 3563
Sequence: 1 MYYQGTSDNTNIQADHQQRH.....GNKRFPTQLLPDFVTGNLGT 688

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10624201/runat_27072005_122957_10905/app_query.fasta_1.839
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10624201_@CGN_1_1_5081_@runat_27072005_122957_10905 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3563	100.0	2735	8	AF406697	AF406697 Solanum t
2	3294.5	92.5	2636	8	AF375966	AF375966 Lycopersi
3	3294.5	92.5	2687	8	BT013459	BT013459 Lycopersi
4	1332	37.4	1898	8	AF406698	AF406698 Solanum t
5	1150.5	32.3	2043	6	AX506642	AX506642 Sequence
6	1150.5	32.3	2043	8	BT008422	BT008422 Arabidops
7	1150.5	32.3	2074	8	AY096548	AY096548 Arabidops
8	1150.5	32.3	2572	8	AY072175	AY072175 Arabidops
9	1150.5	32.3	2578	8	AY099854	AY099854 Arabidops
10	1148.5	32.2	2385	8	AF353094	AF353094 Arabidops
11	1127.5	31.6	2128	8	AF406702	AF406702 Solanum t
12	999	28.0	120938	8	AC007017	AC007017 Arabidops
13	967.5	27.2	2489	8	AK120844	AK120844 Oryza sat
14	946	26.6	2643	8	AK120813	AK120813 Oryza sat
15	945	26.5	3030	8	AK070465	AK070465 Oryza sat
16	931.5	26.1	2065	8	AF406703	AF406703 Solanum t
17	928	26.0	1896	8	AK104872	AK104872 Oryza sat
18	926.5	26.0	2630	8	AK069994	AK069994 Oryza sat
19	922.5	25.9	2523	8	AY524972	AY524972 Oryza sat
20	911.5	25.6	2570	8	AK067939	AK067939 Oryza sat
21	898.5	25.2	2476	8	AK121356	AK121356 Oryza sat
22	891.5	25.0	2952	8	AK070543	AK070543 Oryza sat
23	876	24.6	28712	8	AC145458	AC145458 Cicer ari
24	839.5	23.6	2636	8	AK068034	AK068034 Oryza sat
25	837.5	23.5	2698	8	AK099829	AK099829 Oryza sat
26	835.5	23.4	2414	8	AK105752	AK105752 Oryza sat
27	826	23.2	2050	8	AF353092	AF353092 Arabidops
28	825.5	23.2	1906	8	AY570508	AY570508 Arabidops
29	824	23.1	1884	6	AX506210	AX506210 Sequence
30	824	23.1	1884	8	AY142019	AY142019 Arabidops
31	824	23.1	2124	8	AY058179	AY058179 Arabidops
32	819	23.0	1946	8	AY598452	AY598452 Arabidops
33	819	23.0	1973	8	AY062545	AY062545 Arabidops
c 34	816.5	22.9	146585	8	CNS08CAM	AL831809 Oryza sat
35	809	22.7	1950	8	AF353093	AF353093 Arabidops
36	803	22.5	1824	8	AY085278	AY085278 Arabidops
37	802	22.5	1449	6	AX412914	AX412914 Sequence
38	802	22.5	1599	6	AX507459	AX507459 Sequence
39	802	22.5	1599	8	BT012291	BT012291 Arabidops
40	802	22.5	1634	8	BT000133	BT000133 Arabidops
41	797.5	22.4	2545	8	AF173816	AF173816 Arabidops
42	797.5	22.4	3162	8	AY050459	AY050459 Arabidops
43	797	22.4	1449	8	BT000609	BT000609 Arabidops
44	797	22.4	2126	8	AY056796	AY056796 Arabidops
45	793.5	22.3	3132	8	GGN318871	AJ318871 Gnetum gn

ALIGNMENTS

RESULT 1
AF406697
LOCUS AF406697 2735 bp mRNA linear PLN 14-JUL-2003
DEFINITION Solanum tuberosum BEL1-related homeotic protein 5 (Bel5) mRNA, complete cds.
ACCESSION AF406697
VERSION AF406697.1 GI:22652114
KEYWORDS .
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiids; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 2735)
AUTHORS Chen,H., Rosin,F.M., Prat,S. and Hannapel,D.J.
TITLE Interacting transcription factors from the three-amino acid loop extension superclass regulate tuber formation
JOURNAL Plant Physiol. 132 (3), 1391-1404 (2003)
PUBMED 12857821
REFERENCE 2 (bases 1 to 2735)
AUTHORS Chen,H., Rosin,F.M. and Hannapel,D.J.
TITLE A KNOX protein of potato interacts with several members of the TALE family of transcription factors
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 2735)
AUTHORS Chen,H., Rosin,F.M. and Hannapel,D.J.
TITLE Direct Submission
JOURNAL Submitted (07-AUG-2001) Department of Horticulture, Iowa State

```
          University, 257 Horticulture Hall, Ames, IA 50011-1100, USA
FEATURES             Location/Qualifiers
     source          1. .2735
                     /organism="Solanum tuberosum"
                     /mol_type="mRNA"
                     /cultivar="Desiree"
                     /db_xref="taxon:4113"
     gene            1. .2735
                     /gene="Bel5"
     CDS             148. .2214
                     /gene="Bel5"
                     /note="StBel5"
                     /codon_start=1
                     /product="BEL1-related homeotic protein 5"
                     /protein_id="AAN03621.1"
                     /db_xref="GI:22652115"
                     /translation="MYYQGTSDNTNIQADHQQRHNHGNSNNNNIQTLYLMNPNNYMQG
                     YTTSDTQQQQQLLFLNSSPAASNALCHANIQHAPLQQQHFVGVPLPAVSLHDQINHHG
                     LLQRMWNNQDQSQQVIVPSSTGVSATSCGGITTDLASQLAFQRPIPTPQHRQQQQQQG
                     GLSLSLSPQLQQQISFNNNISSSSPRTNNVTIRGTLDGSSSNMVLGSKYLKAAQELLD
                     EVVNIVGKSIKGDDQKKDNSMNKESMPLASDVNTNSSGGGESSSRQKNEVAVELTTAQ
                     RQELQMKKAKLLAMLEEVEQRYRQYHHQMQIIVLSFEQVAGIGSAKSYTQLALHAISK
                     QFRCLKDAIAEQVKATSKSLGEEEGLGGKIEGSRLKFVDHHLRQQRALQQIGMMQPNA
                     WRPQRGLPERAVSVLRAWLFEHFLHPYPKDSDKIMLAKQTGLTRSQVSNWFINARVRL
                     WKPMVEEMYLEEVKNQEQNSTNTSGDNKNETNISAPNEEKHPIITSSLLQDGITTTQ
                     AEISTSTISTSPTAGASLHHAHNFSFLGSFNMDNTTTTVDHIENNAKKQRNDMHKFSP
                     SSILSSVDMEAKARESSNKGFTNPLMAAYAMGDFGRFDPHDQQMTANFHGNNGVSLTL
                     GLPPSENLAMPVSQQNYLSNDLGSRSEMGSHYNRMGYENIDFQSGNKRFPTQLLPDFV
                     TGNLGT"
ORIGIN

Alignment Scores:
Pred. No.:              1.11e-249       Length:         2735
Score:                  3563.00         Matches:        688
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     8               Gaps:           0

US-10-624-201A-2 (1-688) x AF406697 (1-2735)

Qy          1 MetTyrTyrGlnGlyThrSerAspAsnThrAsnIleGlnAlaAspHisGlnGlnArgHis 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 ATGTACTATCAAGGAACCTCGGATAATACTAATATACAAGCTGATCATCAACAACGTCAT 207

Qy         21 AsnHisGlyAsnSerAsnAsnAsnAsnIleGlnThrLeuTyrLeuMetAsnProAsnAsn 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 AATCATGGGAATAGTAATAATAATAATATTCAGACACTTTATTTGATGAACCCTAACAAT 267

Qy         41 TyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGlnGlnGlnLeuLeuPheLeuAsn 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 TATATGCAAGGCTACACTACTTCTGACACACAGCAGCAGCAGCAGTTACTTTTCCTGAAT 327

Qy         61 SerSerProAlaAlaSerAsnAlaLeuCysHisAlaAsnIleGlnHisAlaProLeuGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 TCTTCACCAGCAGCAAGCAACGCGCTTTGCCATGCGAATATACAACACGCGCCGCTGCAA 387

Qy         81 GlnGlnHisPheValGlyValProLeuProAlaValSerLeuHisAspGlnIleAsnHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 CAGCAGCACTTTGTCGGTGTGCCTCTTCCGGCAGTAAGTTTGCACGATCAGATCAATCAT 447

Qy        101 HisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnSerGlnGlnValIleValPro 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 CATGGACTTTTACAGCGCATGTGGAACAACCAAGATCAATCTCAGCAGGTGATAGTACCA 507

Qy        121 SerSerThrGlyValSerAlaThrSerCysGlyGlyIleThrThrAspLeuAlaSerGln 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 TCGTCGACGGGGGTTTCTGCCACGTCATGTGGCGGGATCACCACGGACTTGGCGTCTCAA 567

Qy        141 LeuAlaPheGlnArgProIleProThrProGlnHisArgGlnGlnGlnGlnGlnGlnGly 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 TTGGCGTTTCAGAGGCCGATTCCGACACCACAACACCGACAGCAGCAACAACAGCAAGGC 627

Qy        161 GlyLeuSerLeuSerLeuSerProGlnLeuGlnGlnGlnIleSerPheAsnAsnAsnIle 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        628 GGTCTATCTCTAAGCCTTTCTCCTCAGCTACAACAGCAAATTAGTTTCAATAACAATATT 687

Qy        181 SerSerSerSerProArgThrAsnAsnValThrIleArgGlyThrLeuAspGlySerSer 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        688 TCATCCTCATCACCAAGGACAAATAATGTTACTATTAGGGGAACATTAGATGGAAGTTCT 747

Qy        201 SerAsnMetValLeuGlySerLysTyrLeuLysAlaAlaGlnGluLeuLeuAspGluVal 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        748 AGCAACATGGTTTTAGGCTCTAAGTATCTGAAAGCTGCACAAGAGCTTCTTGATGAAGTT 807

Qy        221 ValAsnIleValGlyLysSerIleLysGlyAspAspGlnLysLysAspAsnSerMetAsn 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        808 GTTAATATTGTTGGAAAAAGCATCAAAGGAGATGATCAAAAGAAGGATAATTCAATGAAT 867

Qy        241 LysGluSerMetProLeuAlaSerAspValAsnThrAsnSerSerGlyGlyGlyGluSer 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        868 AAAGAATCAATGCCTTTGGCTAGTGATGTCAACACTAATAGTTCTGGTGGTGGTGAAAGT 927

Qy        261 SerSerArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGlnArgGlnGluLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        928 AGCAGCAGGCAGAAAAATGAAGTTGCTGTTGAGCTTACAACTGCTCAAAGACAAGAACTT 987

Qy        281 GlnMetLysLysAlaLysLeuLeuAlaMetLeuGluGluValGluGlnArgTyrArgGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        988 CAAATGAAAAAAGCCAAGCTTCTTGCCATGCTTGAAGAGGTGGAGCAAAGGTACAGACAG 1047

Qy        301 TyrHisHisGlnMetGlnIleIleValLeuSerPheGluGlnValAlaGlyIleGlySer 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1048 TACCATCACCAAATGCAAATAATTGTATTATCATTTGAGCAAGTAGCAGGAATTGGATCA 1107

Qy        321 AlaLysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArgCysLeuLys 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1108 GCCAAATCATACACTCAATTAGCTTTGCATGCAATTTCGAAGCAATTCAGATGCCTAAAG 1167

Qy        341 AspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGluGluGluGlyLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1168 GATGCAATTGCTGAGCAAGTAAAGGCGACGAGCAAGAGTTTAGGTGAAGAGGAAGGCTTG 1227

Qy        361 GlyGlyLysIleGluGlySerArgLeuLysPheValAspHisHisLeuArgGlnGlnArg 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1228 GGAGGGAAAATCGAAGGCTCAAGACTCAAATTTGTGGACCATCATCTAAGGCAACAACGC 1287

Qy        381 AlaLeuGlnGlnIleGlyMetMetGlnProAsnAlaTrpArgProGlnArgGlyLeuPro 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1288 GCGCTGCAACAGATAGGAATGATGCAACCAAATGCTTGGAGACCCCAAAGAGGTTTACCT 1347

Qy        401 GluArgAlaValSerValLeuArgAlaTrpLeuPheGluHisPheLeuHisProTyrPro 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1348 GAAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTTTTCGAGCATTTTCTTCATCCTTACCCA 1407

Qy        421 LysAspSerAspLysIleMetLeuAlaLysGlnThrGlyLeuThrArgSerGlnValSer 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1408 AAGGATTCAGACAAAATCATGCTTGCTAAGCAAACGGGGCTAACAAGGAGCCAGGTGTCT 1467

Qy        441 AsnTrpPheIleAsnAlaArgValArgLeuTrpLysProMetValGluGluMetTyrLeu 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1468 AACTGGTTCATAAATGCTCGAGTTCGATTATGGAAGCCAATGGTAGAAGAAATGTACTTG 1527

Qy        461 GluGluValLysAsnGlnGluGlnAsnSerThrAsnThrSerGlyAspAsnLysAsnLys 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1528 GAAGAAGTGAAGAATCAAGAACAAAACAGTACTAATACTTCAGGAGATAACAAAAACAAA 1587

Qy        481 GluThrAsnIleSerAlaProAsnGluGluLysHisProIleIleThrSerSerLeuLeu 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1588 GAGACCAATATAAGTGCTCCAAATGAAGAGAAACATCCAATTATTACTAGCAGCTTATTA 1647

Qy        501 GlnAspGlyIleThrThrThrGlnAlaGluIleSerThrSerThrIleSerThrSerPro 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1648 CAAGATGGTATTACTACTACTCAAGCAGAAATTTCTACCTCAACTATTTCAACTTCCCCT 1707

Qy        521 ThrAlaGlyAlaSerLeuHisHisAlaHisAsnPheSerPheLeuGlySerPheAsnMet 540
```

```
Db       1708 ACTGCAGGTGCTTCACTTCATCATGCTCACAATTTCTCCTTCCTTGGTTCATTCAACATG 1767

Qy        541 AspAsnThrThrThrThrValAspHisIleGluAsnAsnAlaLysLysGlnArgAsnAsp 560

Db       1768 GATAATACTACTACTACTGTTGATCATATTGAAAACAACGCGAAAAAGCAAAGAAATGAC 1827

Qy        561 MetHisLysPheSerProSerSerIleLeuSerSerValAspMetGluAlaLysAlaArg 580

Db       1828 ATGCACAAGTTTTCTCCAAGTAGTATTCTTTCATCTGTTGACATGGAAGCCAAAGCTAGA 1887

Qy        581 GluSerSerAsnLysGlyPheThrAsnProLeuMetAlaAlaTyrAlaMetGlyAspPhe 600

Db       1888 GAATCATCAAATAAAGGGTTTACTAATCCTTTAATGGCAGCATACGCGATGGGAGATTTT 1947

Qy        601 GlyArgPheAspProHisAspGlnGlnMetThrAlaAsnPheHisGlyAsnAsnGlyVal 620

Db       1948 GGAAGGTTTGATCCTCATGATCAACAAATGACCGCGAATTTTCATGGAAATAATGGTGTC 2007

Qy        621 SerLeuThrLeuGlyLeuProProSerGluAsnLeuAlaMetProValSerGlnGlnAsn 640

Db       2008 TCTCTTACTTTAGGACTTCCTCCTTCTGAAAACCTAGCCATGCCAGTGAGCCAACAAAAT 2067

Qy        641 TyrLeuSerAsnAspLeuGlySerArgSerGluMetGlySerHisTyrAsnArgMetGly 660

Db       2068 TACCTTTCTAATGACTTGGGAAGTAGGTCTGAAATGGGGAGTCATTACAATAGAATGGGA 2127

Qy        661 TyrGluAsnIleAspPheGlnSerGlyAsnLysArgProThrGlnLeuLeuProAsp 680

Db       2128 TATGAAAACATTGATTTTCAGAGTGGGAATAAGCGATTTCCGACTCAACTATTACCAGAT 2187

Qy        681 PheValThrGlyAsnLeuGlyThr 688

Db       2188 TTTGTTACAGGTAATCTAGGAACA 2211
```

RESULT 2
AF375966

```
LOCUS       AF375966               2636 bp    mRNA    linear   PLN 02-JUN-2003
DEFINITION  Lycopersicon esculentum bell-like homeodomain protein 2 (BL2) mRNA,
            complete cds.
ACCESSION   AF375966
VERSION     AF375966.1  GI:31323446
KEYWORDS    .
SOURCE      Lycopersicon esculentum (tomato)
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1  (bases 1 to 2636)
  AUTHORS   Ron,N., Parnis,A. and Lifschitz,E.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-MAY-2001) Biology, Technion, Haifa, Israel
FEATURES             Location/Qualifiers
     source          1. .2636
                     /organism="Lycopersicon esculentum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4081"
     gene            1. .2636
                     /gene="BL2"
     CDS             100. .2199
                     /gene="BL2"
                     /note="TBL2"
                     /codon_start=1
                     /product="bell-like homeodomain protein 2"
                     /protein_id="AAP47025.1"
                     /db_xref="GI:31323447"
                     /translation="MYYQGTSDNNIQADHHQQQHNNLGNSNNNIQTLYLMNPNSYMQG
                     YTTTDTQQHLQQQQNQHQLLFLNSAPAGGNALSHANIQHAPLQQQHFVGVPLPAVSLH
                     DQINHHGLLQRMWNNQDQSQQVIVPSSTVVSATSCGGTTTDLASQLAFQRPIVVSPTP
                     QHRQQQQQQGGLSLSLSPQQQQQISFNNNISSSSPRTNNVTIRGTMDGCSSNMILGSK
                     YLKAAQELLDEVVNIVGKSNKGDDQKKDNSMNKELIPLVSDVNTNSSGGGGGESSSRQ
                     KNEVAIELTTAQRQELQMKKAKLLAMLEEVEQRYRQYHHQMQIIVSSFEQVAGVGSAK
                     SYTQLALHAISKQFRCLKDAISEQVKATSKSLGEDEGLGGKIEGSRLKFVDHHLRQQR
                     ALQQLGMMQPNAWRPQRGLPERAVSVLRAWLFEHFLHPYPKDSDKIMLAKQTGLTRSQ
                     VSNWFINARVRLWKPMVEEMYLEEVKNQEQNSSNTSGDNKNKETNISAPNEEKQPIIT
                     SSLLQDGTTQAEISTSTISTSPTAGASLHHAHNFSFLGSFNMENTTTTVDHIENNAKK
                     PRNHDMHKFSPSSILSSVEMEAKARESTNKGFTNPLMAAYAMGDFGRFDPHDQQMTAN
                     FHGNNGVSLTLGLPPSENLAMPVSQQNYLSNELGSRPEIGSHYNRMGYENIDFQSGNK
                     RFPTQLLPDFVTGNLGT"
ORIGIN
```

Alignment Scores:

Pred. No.:	3.71e-230	Length:	2636
Score:	3294.50	Matches:	647
Percent Similarity:	95.16%	Conservative:	21
Best Local Similarity:	92.17%	Mismatches:	17
Query Match:	92.46%	Indels:	17
DB:	8	Gaps:	7

US-10-624-201A-2 (1-688) x AF375966 (1-2636)

```
Qy          1 MetTyrTyrGlnGlyThrSerAspAsnThrAsnIleGlnAlaAsp---HisGlnGlnArg 19

Db        100 ATGTACTATCAAGGAACCTCGGATAAT---AATATACAAGCTGATCATCATCAACAACAA 156

Qy         20 HisAsnHisGlyAsnSerAsnAsnAsnAsnIleGlnThrLeuTyrLeuMetAsnProAsn 39

Db        157 CATAATAATCTTGGGAATAGTAATAATAATATTCAGACGCTGTATCTGATGAACCCTAAC 216

Qy         40 AsnTyrMetGlnGlyTyrThrThrSerAspThrGlnGln---------------------- 52

Db        217 AGTTATATGCAAGGCTACACTACAACTGACACACAGCAGCACCTGCAACAGCAACAGAAC 276

Qy         53 GlnGlnGlnLeuLeuPheLeuAsnSerSerProAlaAlaSerAsnAlaLeuCysHisAla 72

Db        277 CAGCACCAGTTGCTTTTTCTGAATTCAGCACCAGCAGGAGGCAACGCGCTTAGCCATGCG 336

Qy         73 AsnIleGlnHisAlaProLeuGlnGlnGlnHisPheValGlyValProLeuProAlaVal 92

Db        337 AATATACAACACGCGCCTCTGCAACAGCAGCACTTTGTCGGCGTGCCCCTTCCGGCAGTG 396

Qy         93 SerLeuHisAspGlnIleAsnHisHisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAsp 112

Db        397 AGTTTGCATGATCAGATCAATCATCACGGACTTTTACAGCGTATGTGGAACAATCAAGAC 456

Qy        113 GlnSerGlnGlnValIleValProSerSerThrGlyValSerAlaThrSerCysGlyGly 132

Db        457 CAATCTCAGCAGGTGATAGTACCGTCGTCGACGGTGGTTTCTGCCACGTCATGTGGCGGG 516

Qy        133 IleThrThrAspLeuAlaSerGlnLeuAlaPheGlnArgProIle---------ProThr 149

Db        517 ACCACCACGGACTTGGCGTCTCAACTGGCGTTTCAGAGGCCGATTGTGGTGTCTCCGACA 576

Qy        150 ProGlnHisArgGlnGlnGlnGlnGlnGlyGlyLeuSerLeuSerLeuSerProGln 169

Db        577 CCACAACACCGACAACAACAACAACAGCAAGGGGGTCTATCACTGAGCCTTTCTCCTCAG 636

Qy        170 LeuGlnGlnGlnIleSerPheAsnAsnAsnIleSerSerSerSerProArgThrAsnAsn 189

Db        637 CAACAGCAACAAATTAGTTTCAATAACAATATTTCATCATCATCACCAAGGACAAATAAT 696

Qy        190 ValThrIleArgGlyThrLeuAspGlySerSerSerAsnMetValLeuGlySerLysTyr 209

Db        697 GTTACTATTAGAGGAACAATGGATGGATGTTCTAGCAACATGATTTTAGGCTCTAAGTAT 756

Qy        210 LeuLysAlaAlaGlnGluLeuLeuAspGluValValAsnIleValGlyLysSerIleLys 229

Db        757 CTGAAAGCTGCACAAGAGCTTCTTGATGAAGTGGTTAATATTGTTGGAAAAAGTAACAAA 816

Qy        230 GlyAspAspGlnLysLysAspAsnSerMetAsnLysGluSerMetProLeuAlaSerAsp 249

Db        817 GGAGATGATCAAAAGAAGGATAATTCAATGAATAAAGAATTAATTCCTTTGGTTAGTGAT 876

Qy        250 ValAsnThrAsnSerSer------GlyGlyGlyGluSerSerSerArgGlnLysAsnGlu 267

Db        877 GTCAATACTAATAGTTCTGGTGGTGGTGGTGGTGAAAGTAGCAGCAGGCAGAAAAATGAA 936
```

```
Qy        268 ValAlaValGluLeuThrThrAlaGlnArgGlnGluLeuGlnMetLysLysAlaLysLeu 287
              ||||||:::||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        937 GTTGCCATTGAGCTTACAACTGCTCAAAGACAAGAACTTCAAATGAAAAAAGCCAAGCTT 996

Qy        288 LeuAlaMetLeuGluGluValGluGlnArgTyrArgGlnTyrHisHisGlnMetGlnIle 307
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        997 CTTGCCATGCTTGAAGAGGTGGAGCAAAGGTACAGACAGTACCATCACCAAATGCAAATA 1056

Qy        308 IleValLeuSerPheGluGlnValAlaGlyIleGlySerAlaLysSerTyrThrGlnLeu 327
              ||||||   ||||||||||||||||||:::||||||||||||||||||||||||||||||
Db       1057 ATTGTATCATCATTTGAGCAAGTAGCAGGAGTTGGATCAGCAAAATCATACACTCAATTA 1116

Qy        328 AlaLeuHisAlaIleSerLysGlnPheArgCysLeuLysAspAlaIleAlaGluGlnVal 347
              ||||||||||||||||||||||||||||||||||||||||||:::|||||||||
Db       1117 GCTTTGCATGCAATTTCAAAGCAATTCAGATGCCTAAAGGATGCAATTAGTGAGCAAGTG 1176

Qy        348 LysAlaThrSerLysSerLeuGlyGluGluGluGlyLeuGlyGlyLysIleGluGlySer 367
              ||||||||||||||||||||||||:::|||||||||||||||||||||||||||||||||
Db       1177 AAGGCGACGAGCAAGAGTTTAGGTGAAGATGAAGGCTTGGGAGGGAAAATTGAAGGCTCA 1236

Qy        368 ArgLeuLysPheValAspHisHisLeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMet 387
              ||||||||||||||||||||||||||||||||||||||||||||||||:::|||||||
Db       1237 AGACTCAAATTTGTGGACCATCATCTAAGGCAACAACGCGCGCTGCAACAGCTAGGAATG 1296

Qy        388 MetGlnProAsnAlaTrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeu 407
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1297 ATGCAACCAAATGCTTGGAGACCCCAAAGAGGTTTACCTGAAAGAGCTGTCTCTGTCCTT 1356

Qy        408 ArgAlaTrpLeuPheGluHisPheLeuHisProTyrProLysAspSerAspLysIleMet 427
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1357 CGTGCTTGGCTTTTCGAGCATTTTCTTCATCCTTACCCCAAAGACTCAGACAAAATCATG 1416

Qy        428 LeuAlaLysGlnThrGlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArg 447
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1417 CTTGCTAAGCAAACGGGGCTAACAAGGAGCCAGGTCTCTAACTGGTTCATTAATGCTCGA 1476

Qy        448 ValArgLeuTrpLysProMetValGluGluMetTyrLeuGluGluValLysAsnGlnGlu 467
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1477 GTTCGATTATGGAAGCCAATGGTTGAAGAAATGTACTTGGAAGAAGTGAAGAATCAAGAA 1536

Qy        468 GlnAsnSerThrAsnThrSerGlyAspAsnLysAsnLysGluThrAsnIleSerAlaPro 487
              |||||||||:::||||||||||||||||||||||||||||||||||||||||||||||||
Db       1537 CAAAACAGTAGTAATACTTCAGGAGATAACAAAAACAAAGAGACGAATATAAGTGCTCCA 1596

Qy        488 AsnGluGluLysHisProIleIleThrSerSerLeuLeuGlnAspGlyIleThrThrThr 507
              ||||||||||   ||||||||||||||||||||||||||||||||||      ||||||
Db       1597 AATGAAGAGAAACAACCAATTATTACTAGCAGCTTATTACAAGATGGT------ACTACT 1650

Qy        508 GlnAlaGluIleSerThrSerThrIleSerThrSerProThrAlaGlyAlaSerLeuHis 527
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1651 CAAGCAGAAATTTCTACCTCAACTATTTCAACTTCCCCTACTGCAGGTGCTTCACTTCAT 1710

Qy        528 HisAlaHisAsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrThrVal 547
              ||||||||||||||||||||||||||||||||||||:::||||||||||||||||||||
Db       1711 CATGCTCACAACTTCTCCTTCCTTGGTTCATTCAACATGGAGAATACTACTACTACTGTT 1770

Qy        548 AspHisIleGluAsnAsnAlaLysLysGlnArgAsn---AspMetHisLysPheSerPro 566
              ||||||||||||||||||||||||||||||||||||   ||||||  ||||||||||||||
Db       1771 GATCATATTGAAAACAACGCGAAAAAGCCAAGAAATCATGACATGCACAAGTTTTCTCCA 1830

Qy        567 SerSerIleLeuSerSerValAspMetGluAlaLysAlaArgGluSerSerAsnLysGly 586
              ||||||||||||||||||||:::||||||||||||||||||||:::||||||||||
Db       1831 AGTAGTATTCTTTCATCTGTTGAAATGGAAGCCAAAGCTAGAGAATCAACAAACAAAGGG 1890

Qy        587 PheThrAsnProLeuMetAlaAlaTyrAlaMetGlyAspPheGlyArgPheAspProHis 606
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1891 TTTACTAATCCTTTAATGGCAGCGTACGCGATGGGAGATTTTGGAAGGTTTGATCCTCAT 1950

Qy        607 AspGlnGlnMetThrAlaAsnPheHisGlyAsnAsnGlyValSerLeuThrLeuGlyLeu 626
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1951 GATCAACAAATGACCGCGAATTTTCATGGAAATAATGGTGTATCTCTTACTTTAGGACTT 2010

Qy        627 ProProSerGluAsnLeuAlaMetProValSerGlnGlnAsnTyrLeuSerAsnAspLeu 646
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db       2011 CCTCCTTCTGAAAACCTAGCCATGCCAGTGAGCCAACAAAATTATCTTTCTAATGAGTTG 2070

Qy        647 GlySerArgSerGluMetGlySerHisTyrAsnArgMetGlyTyrGluAsnIleAspPhe 666
              |||||||||   |||:::||||||||||||||||||||||||||||||||||||||||||
Db       2071 GGAAGTAGGCCTGAAATAGGAAGTCATTACAATAGAATGGGATATGAAAACATTGATTTT 2130

Qy        667 GlnSerGlyAsnLysArgPheProThrGlnLeuLeuProAspPheValThrGlyAsnLeu 686
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2131 CAAAGTGGGAATAAGCGATTTCCAACTCAACTATTGCCAGATTTTGTTACAGGTAATCTT 2190

Qy        687 GlyThr 688
              ||||||
Db       2191 GGAACA 2196

RESULT 3
BT013459
LOCUS       BT013459               2687 bp    mRNA    linear   PLN 11-MAY-2004
DEFINITION  Lycopersicon esculentum clone 132117F, mRNA sequence.
ACCESSION   BT013459
VERSION     BT013459.1  GI:47104874
KEYWORDS    FLI_CDNA.
SOURCE      Lycopersicon esculentum (tomato)
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1  (bases 1 to 2687)
  AUTHORS   Kirkness,E.F., Wang,W. and Vazeille,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-MAY-2004) The Institute for Genomic Research, 9712
            Medical Center Drive, Rockville, MD 20850, USA
FEATURES             Location/Qualifiers
     source          1. .2687
                     /organism="Lycopersicon esculentum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4081"
                     /clone="132117F"
                     /tissue_type="fruit"
                     /note="TZXBC46"
ORIGIN

Alignment Scores:
Pred. No.:              3.8e-230        Length:         2687
Score:                  3294.50         Matches:        647
Percent Similarity:     95.16%          Conservative:   21
Best Local Similarity:  92.17%          Mismatches:     17
Query Match:            92.46%          Indels:         17
DB:                     8               Gaps:           7

US-10-624-201A-2 (1-688) x BT013459 (1-2687)

Qy          1 MetTyrTyrGlnGlyThrSerAspAsnThrAsnIleGlnAlaAsp---HisGlnGlnArg 19
              ||||||||||||||||||||||||||   ||||||||||||||   |||||||||:::
Db        154 ATGTACTATCAAGGAACCTCGGATAAT---AATATACAAGCTGATCATCATCAACAACAA 210

Qy         20 HisAsnHisGlyAsnSerAsnAsnAsnAsnIleGlnThrLeuTyrLeuMetAsnProAsn 39
              ||||||:::      ::::::|||||||||||||||||||||||||||||||||||||
Db        211 CATAATAATCTTGGGAATAGTAATAATAATATTCAGACGCTGTATCTGATGAACCCTAAC 270

Qy         40 AsnTyrMetGlnGlyTyrThrThrSerAspThrGlnGln---------------------- 52
              :::|||||||||||||||||||||:::|||||||||||
Db        271 AGTTATATGCAAGGCTACACTACAACTGACACACAGCAGCACCTGCAACAGCAACAGAAC 330

Qy         53 GlnGlnGlnLeuLeuPheLeuAsnSerSerProAlaAlaSerAsnAlaLeuCysHisAla 72
              |||   ||||||||||||||||||:::||||||      |||||||||   ||||||
Db        331 CAGCACCAGTTGCTTTTTCTGAATTCAGCACCAGCAGGAGGCAACGCGCTTAGCCATGCG 390

Qy         73 AsnIleGlnHisAlaProLeuGlnGlnGlnHisPheValGlyValProLeuProAlaVal 92
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 AATATACAACACGCGCCTCTGCAACAGCAGCACTTTGTCGGCGTGCCCCTTCCGGCAGTG 450
```

```
Qy        93 SerLeuHisAspGlnIleAsnHisHisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAsp 112
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       451 AGTTTGCATGATCAGATCAATCATCACGGACTTTTACAGCGTATGTGGAACAATCAAGAC 510

Qy       113 GlnSerGlnGlnValIleValProSerSerThrGlyValSerAlaThrSerCysGlyGly 132
             ||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||
Db       511 CAATCTCAGCAGGTGATAGTACCGTCGTCGACGGTGGTTTCTGCCACGTCATGTGGCGGG 570

Qy       133 IleThrThrAspLeuAlaSerGlnLeuAlaPheGlnArgProIle---------ProThr 149
             ||||||||||||||||||||||||||||||||||||||||||||||         ||||||
Db       571 ACCACCACGGACTTGGCGTCTCAACTGGCGTTTCAGAGGCCGATTGTGGTGTCTCCGACA 630

Qy       150 ProGlnHisArgGlnGlnGlnGlnGlnGlyGlyLeuSerLeuSerLeuSerProGlnGln 169
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       631 CCACAACACCGACAACAACAACAACAGCAAGGCGGTCTATCACTGAGCCTTTCTCCTCAG 690

Qy       170 LeuGlnGlnGlnIleSerPheAsnAsnAsnIleSerSerSerSerProArgThrAsnAsn 189
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       691 CAACAGCAACAAATTAGTTTCAATAACAATATTTCATCATCATCACCAAGGACAAATAAT 750

Qy       190 ValThrIleArgGlyThrLeuAspGlySerSerSerAsnMetValLeuGlySerLysTyr 209
             ||||||||||||||||||:::||||||  ||||||||||||:::||||||||||||||||||
Db       751 GTTACTATTAGAGGAACAATGGATGGATGTTCTAGCAACATGATTTTAGGCTCTAAGTAT 810

Qy       210 LeuLysAlaAlaGlnGluLeuLeuAspGluValValAsnIleValGlyLysSerIleLys 229
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       811 CTGAAAGCTGCACAAGAGCTTCTTGATGAAGTGGTTAATATTGTTGGAAAAAGTAACAAA 870

Qy       230 GlyAspAspGlnLysLysAspAsnSerMetAsnLysGluSerMetProLeuAlaSerAsp 249
             ||||||||||||||||||||||||||||||||||||||||||  :::||||||  ||||||
Db       871 GGAGATGATCAAAAGAAGGATAATTCAATGAATAAAGAATTAATTCCTTTGGTTAGTGAT 930

Qy       250 ValAsnThrAsnSerSer------GlyGlyGlyGluSerSerSerArgGlnLysAsnGlu 267
             ||||||||||||||||||      ||||||||||||||||||||||||||||||||||||||||
Db       931 GTCAATACTAATAGTTCTGGTGGTGGTGGTGGTGAAAGTAGCAGCAGGCAGAAAAATGAA 990

Qy       268 ValAlaValGluLeuThrThrAlaGlnArgGlnGluLeuGlnMetLysLysAlaLysLeu 287
             ||||||:::||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       991 GTTGCCATTGAGCTTACAACTGCTCAAAGACAAGAACTTCAAATGAAAAAAGCCAAGCTT 1050

Qy       288 LeuAlaMetLeuGluGluValGluGlnArgTyrArgGlnTyrHisHisGlnMetGlnIle 307
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1051 CTTGCCATGCTTGAAGAGGTGGAGCAAAGGTACAGACAGTACCATCACCAAATGCAAATA 1110

Qy       308 IleValLeuSerPheGluGlnValAlaGlyIleGlySerAlaLysSerTyrThrGlnLeu 327
             ||||||   ||||||||||||||||||:::||||||||||||||||||||||||||||||||||
Db      1111 ATTGTATCATCATTTGAGCAAGTAGCAGGAGTTGGATCAGCAAAATCATACACTCAATTA 1170

Qy       328 AlaLeuHisAlaIleSerLysGlnPheArgCysLeuLysAspAlaIleAlaGluGlnVal 347
             ||||||||||||||||||||||||||||||||||||||||||||||||:::||||||||||
Db      1171 GCTTTGCATGCAATTTCAAAGCAATTCAGATGCCTAAAGGATGCAATTAGTGAGCAAGTG 1230

Qy       348 LysAlaThrSerLysSerLeuGlyGluGluGluGlyLeuGlyGlyLysIleGluGlySer 367
             ||||||||||||||||||||||||||||||::||||||||||||||||||||||||||||||||
Db      1231 AAGGCGACGAGCAAGAGTTTAGGTGAAGATGAAGGCTTGGGAGGGAAAATTGAAGGCTCA 1290

Qy       368 ArgLeuLysPheValAspHisHisLeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMet 387
             ||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||||||
Db      1291 AGACTCAAATTTGTGGACCATCATCTAAGGCAACAACGCGCGCTGCAACAGCTAGGAATG 1350

Qy       388 MetGlnProAsnAlaTrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeu 407
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1351 ATGCAACCAAATGCTTGGAGACCCCAAAGAGGTTTACCTGAAAGAGCTGTCTCTGTCCTT 1410

Qy       408 ArgAlaTrpLeuPheGluHisPheLeuHisProTyrProLysAspSerAspLysIleMet 427
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1411 CGTGCTTGGCTTTTCGAGCATTTTCTTCATCCTTACCCCAAAGACTCAGACAAAATCATG 1470

Qy       428 LeuAlaLysGlnThrGlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArg 447
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1471 CTTGCTAAGCAAACGGGGCTAACAAGGAGCCAGGTCTCTAACTGGTTCATTAATGCTCGA 1530

Qy       448 ValArgLeuTrpLysProMetValGluGluMetTyrLeuGluGluValLysAsnGlnGlu 467
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1531 GTTCGATTATGGAAGCCAATGGTTGAAGAAATGTACTTGGAAGAAGTGAAGAATCAAGAA 1590

Qy       468 GlnAsnSerThrAsnThrSerGlyAspAsnLysAsnLysGluThrAsnIleSerAlaPro 487
             |||||||||:::||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1591 CAAAACAGTAGTAATACTTCAGGAGATAACAAAAACAAAGAGACGAATATAAGTGCTCCA 1650

Qy       488 AsnGluGluLysHisProIleIleThrSerSerLeuLeuGlnAspGlyIleThrThrThr 507
             ||||||||||||  ||||||||||||||||||||||||||||||||  ||||||
Db      1651 AATGAAGAGAAACAACCAATTATTACTAGCAGCTTATTACAAGATGGT------ACTACT 1704

Qy       508 GlnAlaGluIleSerThrSerThrIleSerThrSerProThrAlaGlyAlaSerLeuHis 527
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1705 CAAGCAGAAATTTCTACCTCAACTATTTCAACTTCCCCTACTGCAGGTGCTTCACTTCAT 1764

Qy       528 HisAlaHisAsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrThrVal 547
             ||||||||||||||||||||||||||||||||||||||||||:::||||||||||||||||||
Db      1765 CATGCTCACAACTTCTCCTTCCTTGGTTCATTCAACATGGAGAATACTACTACTACTGTT 1824

Qy       548 AspHisIleGluAsnAsnAlaLysLysGlnArgAsn---AspMetHisLysPheSerPro 566
             ||||||||||||||||||||||||||||||||||||  ||||||  ||||||||||||||||||
Db      1825 GATCATATTGAAAACAACGCGAAAAAGCCAAGAAATCATGACATGCACAAGTTTTCTCCA 1884

Qy       567 SerSerIleLeuSerSerValAspMetGluAlaLysAlaArgGluSerSerAsnLysGly 586
             ||||||||||||||||||||||||||||||:::||||||||||||||||||:::||||||||||
Db      1885 AGTAGTATTCTTTCATCTGTTGAAATGGAAGCCAAAGCTAGAGAATCAACAAACAAAGGG 1944

Qy       587 PheThrAsnProLeuMetAlaAlaTyrAlaMetGlyAspPheGlyArgPheAspProHis 606
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1945 TTTACTAATCCTTTAATGGCAGCGTACGCGATGGGAGATTTTGGAAGGTTTGATCCTCAT 2004

Qy       607 AspGlnGlnMetThrAlaAsnPheHisGlyAsnAsnGlyValSerLeuThrLeuGlyLeu 626
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2005 GATCAACAAATGACCGCGAATTTTCATGGAAATAATGGTGTATCTCTTACTTTAGGACTT 2064

Qy       627 ProProSerGluAsnLeuAlaMetProValSerGlnGlnAsnTyrLeuSerAsnAspLeu 646
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||||
Db      2065 CCTCCTTCTGAAAACCTAGCCATGCCAGTGAGCCAACAAAATTATCTTTCTAATGAGTTG 2124

Qy       647 GlySerArgSerGluMetGlySerHisTyrAsnArgMetGlyTyrGluAsnIleAspPhe 666
             |||||||||   |||:::||||||||||||||||||||||||||||||||||||||||||||
Db      2125 GGAAGTAGGCCTGAAATAGGAAGTCATTACAATAGAATGGGATATGAAAACATTGATTTT 2184

Qy       667 GlnSerGlyAsnLysArgProProThrGlnLeuLeuProAspPheValThrGlyAsnLeu 686
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2185 CAAAGTGGGAATAAGCGATTTCCAACTCAACTATTGCCAGATTTTGTTACAGGTAATCTT 2244

Qy       687 GlyThr 688
             ||||||
Db      2245 GGAACA 2250

RESULT 4
AF406698
LOCUS       AF406698               1898 bp    mRNA    linear   PLN 14-JUL-2003
DEFINITION  Solanum tuberosum BEL1-related homeotic protein 11 (Bel11) mRNA,
            partial cds.
ACCESSION   AF406698
VERSION     AF406698.1  GI:22652116
KEYWORDS    .
SOURCE      Solanum tuberosum (potato)
  ORGANISM  Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum.
REFERENCE   1  (bases 1 to 1898)
  AUTHORS   Chen,H., Rosin,F.M., Prat,S. and Hannapel,D.J.
  TITLE     Interacting transcription factors from the three-amino acid loop
            extension superclass regulate tuber formation
  JOURNAL   Plant Physiol. 132 (3), 1391-1404 (2003)
   PUBMED   12857821
REFERENCE   2  (bases 1 to 1898)
  AUTHORS   Chen,H., Rosin,F.M. and Hannapel,D.J.
  TITLE     A KNOX protein of potato interacts with several members of the TALE
```

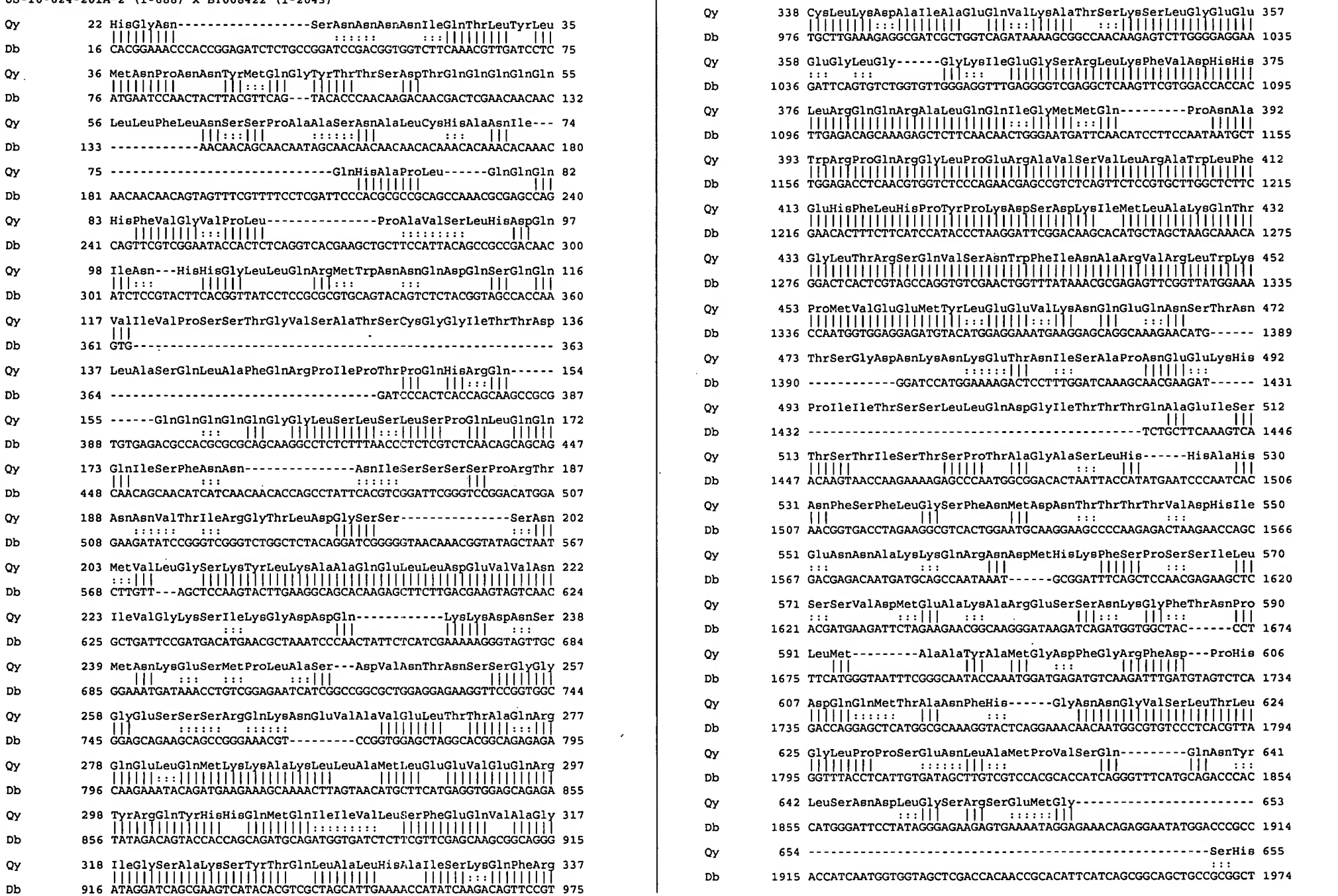
```
          family of transcription factors
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 1898)
  AUTHORS   Chen,H., Rosin,F.M. and Hannapel,D.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-AUG-2001) Department of Horticulture, Iowa State
            University, 257 Horticulture Hall, Ames, IA 50011-1100, USA
FEATURES             Location/Qualifiers
     source          1. .1898
                     /organism="Solanum tuberosum"
                     /mol_type="mRNA"
                     /cultivar="Desiree"
                     /db_xref="taxon:4113"
     gene            <1. .1898
                     /gene="Bel11"
     CDS             <1. .1608
                     /gene="Bel11"
                     /note="StBel11"
                     /codon_start=1
                     /product="BEL1-related homeotic protein 11"
                     /protein_id="AAN03622.1"
                     /db_xref="GI:22652117"
                     /translation="MTFRSSLPLDLREISTTNHQVGILSSSPLPSPGTNTNNINHTRG
                     LGASSSFSISNGMILGSKYLKVAQDLLDEVVNVGKNIKLSDGLESGAKEKHKLDNELI
                     SLASDDVESSSQKNSGVELTTAQRQELQMKKAKLVSMLDEVDQRYRQYHHQMQMIATS
                     FEQTTGIGSSKSYTQLALHTISKQFRCLKDAISGQIKDTSKTLGEEENIGGKIEGSKL
                     KFVDHHLRQQRALQQLGMMQTNAWKPQRGLPERAVSVLRAWLFEHFLHPYPKDSDKII
                     LAKQTGLTRSQVSNWFINARVRLWKPMVEEMYMEEVKKNNQEQNIEPNNNEIVGSKSS
                     VPQEKLPISSNIIHNASPNDISTSTISTSPTGGGGSIPTQTVAGFSFIRSLNMENIDD
                     QRNNKKARNEMQNCSTSTILSMEREIINKVVQDETIKSEKFNNTQTRECYSLMTPNYT
                     MDDQFGTRFNNQNHEQLATTTTFHQGNGHVSLTLGLPPNSENQHNYIGLENHYNQPTH
                     HPNISYENIDFQSGKRYATQLLQDFVS"
ORIGIN

Alignment Scores:
Pred. No.:              1.74e-87        Length:         1898
Score:                  1332.00         Matches:        319
Percent Similarity:     66.84%          Conservative:   56
Best Local Similarity:  56.86%          Mismatches:     100
Query Match:            37.38%          Indels:         86
DB:                     8               Gaps:           27

US-10-624-201A-2 (1-688) x AF406698 (1-1898)

Qy        172 GlnGlnIleSerPheAsnAsnAsn---------IleSerSerSerSerPro--------- 185
Db         34 CGTGAAATTTCAACAACAAATCATCAAGTTGGAATACTATCATCATCACCATTACCATCA 93

Qy        186 ---------ArgThrAsnAsnValThrIleArgGlyThrLeuAspGlySerSerSer--- 201
Db         94 CCAGGAACAAATACCAATAATATCAATCATACTCGAGGATTAGGGGCATCATCATCTTTT 153

Qy        202 ------------AsnMetValLeuGlySerLysTyrLeuLysAlaAlaGlnGluLeuLeu 217
Db        154 TCGATTTCTAATGGGATGATATTGGGTTCTAAGTACCTAAAAGTTGCACAAGATCTTCTT 213

Qy        218 AspGluValValAsnIleValGlyLysSerIleLysGlyAspAsp-------------- 232
Db        214 GATGAAGTTGTTAAT---GTTGGAAAAAACATCAAATTATCAGATGGCTTAGAGAGTGGT 270

Qy        233 GlnLysLysAspAsnSerMetAsnLysGluSerMetProLeuAlaSerAspValAsnThr 252
Db        271 GCAAAGGAGAAACACAAATTGGACAATGAATTAATATCTTTGGCTAGTGAT--------- 321

Qy        253 AsnSerSerGlyGlyGlyGluSerSerSerArgGlnLysAsnGluValAlaValGluLeu 272
Db        322 --------------GATGTTGAAAGCAGCAGCCAAAAAAATAGT---GGTGTTGAACTT 363

Qy        273 ThrThrAlaGlnArgGlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeuGlu 292
Db        364 ACAACAGCTCAAAGACAAGAACTTCAAATGAAGAAAGCCAAGCTTGTTAGCATGCTTGAT 423

Qy        293 GluValGluGlnArgTyrArgGlnTyrHisHisGlnMetGlnIleIleValLeuSerPhe 312
Db        424 GAGGTGGATCAAAGGTATAGACAATACCATCACCAAATGCAAATGATTGCAACATCATTT 483

Qy        313 GluGlnValAlaGlyIleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaIle 332
Db        484 GAGCAAACAACAGGAATTGGATCATCAAAATCATACACACAACTTGCTTTGCACACAATT 543

Qy        333 SerLysGlnPheArgCysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLys 352
Db        544 TCAAAGCAATTTAGATGTTTAAAAGATGCAATTTCTGGGCAAATAAAGGACACTAGCAAA 603

Qy        353 SerLeuGlyGluGluGluGlyLeuGlyGlyLysIleGluGlySerArgLeuLysPheVal 372
Db        604 ACTTTAGGGGAAGAAGAAAACATTGGAGGCAAAATTGAAGGATCAAAGTTGAAATTTGTG 663

Qy        373 AspHisHisLeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGlnProAsnAla 392
Db        664 GATCATCATTTACGCCAACAACGTGCACTACAACAATTAGGGATGATGCAAACCAATGCA 723

Qy        393 TrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeuPhe 412
Db        724 TGGAAGCCTCAAAGAGGTTTGCCAGAAAGAGCGGTTTCAGTTCTCCGCGCTTGGCTTTTC 783

Qy        413 GluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGlnThr 432
Db        784 GAGCATTTTCTTCATCCGTATCCCAAAGATTCAGATAAAATCATCCTTGCTAAGCAAACA 843

Qy        433 GlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLys 452
Db        844 GGGCTAACAAGGAGCCAGGTATCAAATTGGTTTATAAATGCTAGAGTTAGACTATGGAAG 903

Qy        453 ProMetValGluGluMetTyrLeuGluGluValLys------AsnGlnGluGlnAsnSer 470
Db        904 CCAATGGTAGAAGAAATGTACATGGAAGAAGTGAAGAAAAACAATCAAGAACAAAATATT 963

Qy        471 ThrAsnThrSerGlyAspAsnLysAsnLysGluThrAsnIleSerAlaProAsnGluGlu 490
Db        964 GAGCCTAATAACAATGAAATTGTTGGCTCAAAATCAAGTGTTCCA---------CAAGAG 1014

Qy        491 LysHisProIleIleThrSerSerLeuLeuGlnAspGlyIleThrThrThrGlnAlaGlu 510
Db       1015 AAATTACCA---ATTAGTAGCAATATTATTCATAAT---------GCTTCTCCAAATGAT 1062

Qy        511 IleSerThrSerThrIleSerThrSerProThrAlaGlyAlaSer---------LeuHis 527
Db       1063 ATTTCTACTTCCACCATTTCAACATCTCCGACGGGTGGCGGCGGTTCGATTCCGACTCAG 1122

Qy        528 HisAlaHisAsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrThrVal 547
Db       1123 ACGGTTGCAGGTTTCTCCTTCATTAGGTCATTAAACATGGAGAAC------------ATT 1170

Qy        548 AspHisIleGluAsnAsnAlaLysLysGlnArgAsnAspMetHisLysPheSerProSer 567
Db       1171 GATGATCAAAGGAACAAC---AAAAAGGCAAGAAATGAGATGCAAAATTGTTCAACTAGT 1227

Qy        568 SerIleLeuSerSerValAspMetGluAlaLysAlaArgGluSerSerAsnLys------ 585
Db       1228 ACTATTCTCTCA---------ATGGAA---------AGAGAAATCATAAATAAAGTTGTG 1269

Qy        586 ----------------------GlyPheThrAsnPro---------------------- 590
Db       1270 CAAGATGAGACAATCAAAAGTGAAAAGTTCAACAACACACAAACAAGAGAATGTTACTCT 1329

Qy        591 ---LeuMetAlaAlaTyrAlaMetGlyAsp---PheGly---ArgPheAsp------Pro 605
Db       1330 CTAATGACTCCAAATTACACAATGGATGATCAATTTGGAACAAGGTTCAATAATCAAAAT 1389

Qy        606 HisAspGln---GlnMetThrAlaAsnPheHis---GlyAsnAsnGlyValSerLeuThr 623
Db       1390 CATGAACAATTGGCAACAACAACAACTTTTCATCAAGGAAATGGTCATGTTTCTCTTACT 1449

Qy        624 LeuGlyLeuProPro---SerGluAsnLeuAlaMetProValSerGlnGlnAsnTyrLeu 642
```

```
Db       1450 TTAGGGCTTCCACCAAATTCTGAAAAC-----------------CAACACAATTACATT 1491

Qy        643 SerAsnAspLeuGlySerArgSerGluMetGlySerHisTyrAsnArgMetGlyTyrGlu 662
                  |||  :::                :::|||:::      :::   ||||||
Db       1492 GGA------TTGGAAAATCATTACAATCAACCTACACATCATCCAAATATTAGCTATGAA 1545

Qy        663 AsnIleAspPheGlnSerGlyAsnLysArgPheProThrGlnLeuLeuProAspPheVal 682
              ||||||||||||||||||||   ||||||:::   ||||||||||||   |||||||||
Db       1546 AACATTGATTTTCAGAGTGGA---AAGCGATACGCCACTCAACTATTACAAGATTTTGTT 1602

Qy        683 Thr 683
              :::
Db       1603 TCT 1605

RESULT 5
AX506642
LOCUS       AX506642               2043 bp    DNA     linear   PAT 27-SEP-2002
DEFINITION  Sequence 1337 from Patent WO0216655.
ACCESSION   AX506642
VERSION     AX506642.1  GI:23387879
KEYWORDS    .
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Harper,J.F., Kreps,J., Wang,X. and Zhu,T.
  TITLE     Stress-regulated genes of plants, transgenic plants containing
            same, and methods of use
  JOURNAL   Patent: WO 0216655-A 1337 28-FEB-2002;
            The Scripps Research Institute (US) ; Syngenta Participations AG
            (CH)
FEATURES             Location/Qualifiers
     source          1. .2043
                     /organism="Arabidopsis thaliana"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:3702"
ORIGIN

Alignment Scores:
Pred. No.:                3.11e-74       Length:          2043
Score:                    1150.50        Matches:         306
Percent Similarity:       51.81%         Conservative:    81
Best Local Similarity:    40.96%         Mismatches:      201
Query Match:              32.29%         Indels:          159
DB:                       6              Gaps:            28

US-10-624-201A-2 (1-688) x AX506642 (1-2043)

Qy          22 HisGlyAsn-----------------SerAsnAsnAsnAsnIleGlnThrLeuTyrLeu 35
               |||||||||                ::::::      :::|||||||||   |||
Db          16 CACGGAAACCCACCGGAGATCTCTGCCGGATCCGACGGTGGTCTTCAAACGTTGATCCTC 75

Qy          36 MetAsnProAsnAsnTyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGlnGlnGln 55
               |||||||||      |||:::|||   ||||||      |||
Db          76 ATGAATCCAACTACTTACGTTCAG---TACACCCAACAAGACAACGACTCGAACAACAAC 132

Qy          56 LeuLeuPheLeuAsnSerSerProAlaAlaSerAsnAlaLeuCysHisAlaAsnIle--- 74
                     |||:::|||      ::::::|||         :::   |||
Db         133 ------------AACAACAGCAACAATAGCAACAACAACAACACAAACACAAACACAAAC 180

Qy          75 ------------------------------GlnHisAlaProLeu------GlnGlnGln 82
                                            |||||||||              |||
Db         181 AACAACAACAGTAGTTTCGTTTTCCTCGATTCCCACGCGCCGCAGCCAAACGCGAGCCAG 240

Qy          83 HisPheValGlyValProLeu--------------ProAlaValSerLeuHisAspGln 97
               |||||||||:::||||||                :::::::::      |||
Db         241 CAGTTCGTCGGAATACCACTCTCAGGTCACGAAGCTGCTTCCATTACAGCCGCCGACAAC 300

Qy          98 IleAsn---HisHisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnSerGlnGln 116
               |||:::      ||||||          |||:::      :::          |||   |||
Db         301 ATCTCCGTACTTCACGGTTATCCTCCGCGCGTGCAGTACAGTCTCTACGGTAGCCACCAA 360

Qy         117 ValIleValProSerSerThrGlyValSerAlaThrSerCysGlyGlyIleThrThrAsp 136
               |||
Db         361 GTG------------------------------------------------------- 363

Qy         137 LeuAlaSerGlnLeuAlaPheGlnArgProIleProThrProGlnHisArgGln------ 154
                                                        |||   |||:::|||
Db         364 ----------------------------------GATCCCACTCACCAGCAAGCCGCG 387

Qy         155 ------GlnGlnGlnGlnGlnGlyGlyLeuSerLeuSerLeuSerProGlnLeuGlnGln 172
                     :::   |||   ||||||||||||:::||||||   |||   ||||||
Db         388 TGTGAGACGCCACGCGCGCAGCAAGGCCTCTCTTTAACCCTCTCGTCTCAACAGCAGCAG 447

Qy         173 GlnIleSerPheAsnAsn--------------AsnIleSerSerSerSerProArgThr 187
               |||         :::                  ::::::         |||
Db         448 CAACAGCAACATCATCAACAACACCAGCCTATTCACGTCGGATTCGGGTCCGGACATGGA 507

Qy         188 AsnAsnValThrIleArgGlyThrLeuAspGlySerSer--------------SerAsn 202
                  ::::::  :::                 ||||||                :::|||
Db         508 GAAGATATCCGGGTCGGGTCTGGCTCTACAGGATCGGGGGTAACAAACGGTATAGCTAAT 567

Qy         203 MetValLeuGlySerLysTyrLeuLysAlaAlaGlnGluLeuLeuAspGluValValAsn 222
               :::|||     ||||||||||||||||||||||||||||||||||||||||||||||||
Db         568 CTTGTT---AGCTCCAAGTACTTGAAGGCAGCACAAGAGCTTCTTGACGAAGTAGTCAAC 624

Qy         223 IleValGlyLysSerIleLysGlyAspAspGln-----------LysLysAspAsnSer 238
                        :::           |||           ||||||   :::
Db         625 GCTGATTCCGATGACATGAACGCTAAATCCCAACTATTCTCATCGAAAAAGGGTAGTTGC 684

Qy         239 MetAsnLysGluSerMetProLeuAlaSer---AspValAsnThrAsnSerSerGlyGly 257
                  |||  :::  :::      :::|||                        |||||||||
Db         685 GGAAATGATAAACCTGTCGGAGAATCATCGGCCGGCGCTGGAGGAGAAGGTTCCGGTGGC 744

Qy         258 GlyGluSerSerSerArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGlnArg 277
               |||       ::::::  ::::::          |||||||||   ||||||:::|||
Db         745 GGAGCAGAAGCAGCCGGGAAACGT---------CCGGTGGAGCTAGGCACGGCAGAGAGA 795

Qy         278 GlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeuGluGluValGluGlnArg 297
               ||||||:::||||||||||||||||||||       ||||||   ||||||||||||||||
Db         796 CAAGAAATACAGATGAAGAAAGCAAAACTTAGTAACATGCTTCATGAGGTGGAGCAGAGA 855

Qy         298 TyrArgGlnTyrHisHisGlnMetGlnIleIleValLeuSerPheGluGlnValAlaGly 317
               ||||||||||||||  |||||||||:::::::::   ||||||||||   ||||||
Db         856 TATAGACAGTACCACCAGCAGATGCAGATGGTGATCTCTTCGTTCGAGCAAGCGGCAGGG 915

Qy         318 IleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArg 337
               |||||||||||||||||||||||  |||||||||      ||||||:::|||||||||
Db         916 ATAGGATCAGCGAAGTCATACACGTCGCTAGCATTGAAAACCATATCAAGACAGTTCCGT 975

Qy         338 CysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGluGlu 357
               |||||||||:::|||||||||   |||:::||||||   :::|||||||||||||||||||
Db         976 TGCTTGAAAGAGGCGATCGCTGGTCAGATAAAAGCGGCCAACAAGAGTCTTGGGGAGGAA 1035

Qy         358 GluGlyLeuGly------GlyLysIleGluGlySerArgLeuLysPheValAspHisHis 375
               :::   :::          |||:::  ||||||||||||||||||||||||||||||||
Db        1036 GATTCAGTGTCTGGTGTTGGGAGGTTTGAGGGGTCGAGGCTCAAGTTCGTGGACCACCAC 1095

Qy         376 LeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGln---------ProAsnAla 392
               ||||||||||||||||||||||||||:::||||||:::|||         ||||||
Db        1096 TTGAGACAGCAAAGAGCTCTTCAACAACTGGGAATGATTCAACATCCTTCCAATAATGCT 1155

Qy         393 TrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeuPhe 412
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1156 TGGAGACCTCAACGTGGTCTCCCAGAACGAGCCGTCTCAGTTCTCCGTGCTTGGCTCTTC 1215

Qy         413 GluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGlnThr 432
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1216 GAACACTTTCTTCATCCATACCCTAAGGATTCGGACAAGCACATGCTAGCTAAGCAAACA 1275

Qy         433 GlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLys 452
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1276 GGACTCACTCGTAGCCAGGTGTCGAACTGGTTTATAAACGCGAGAGTTCGGTTATGGAAA 1335
```

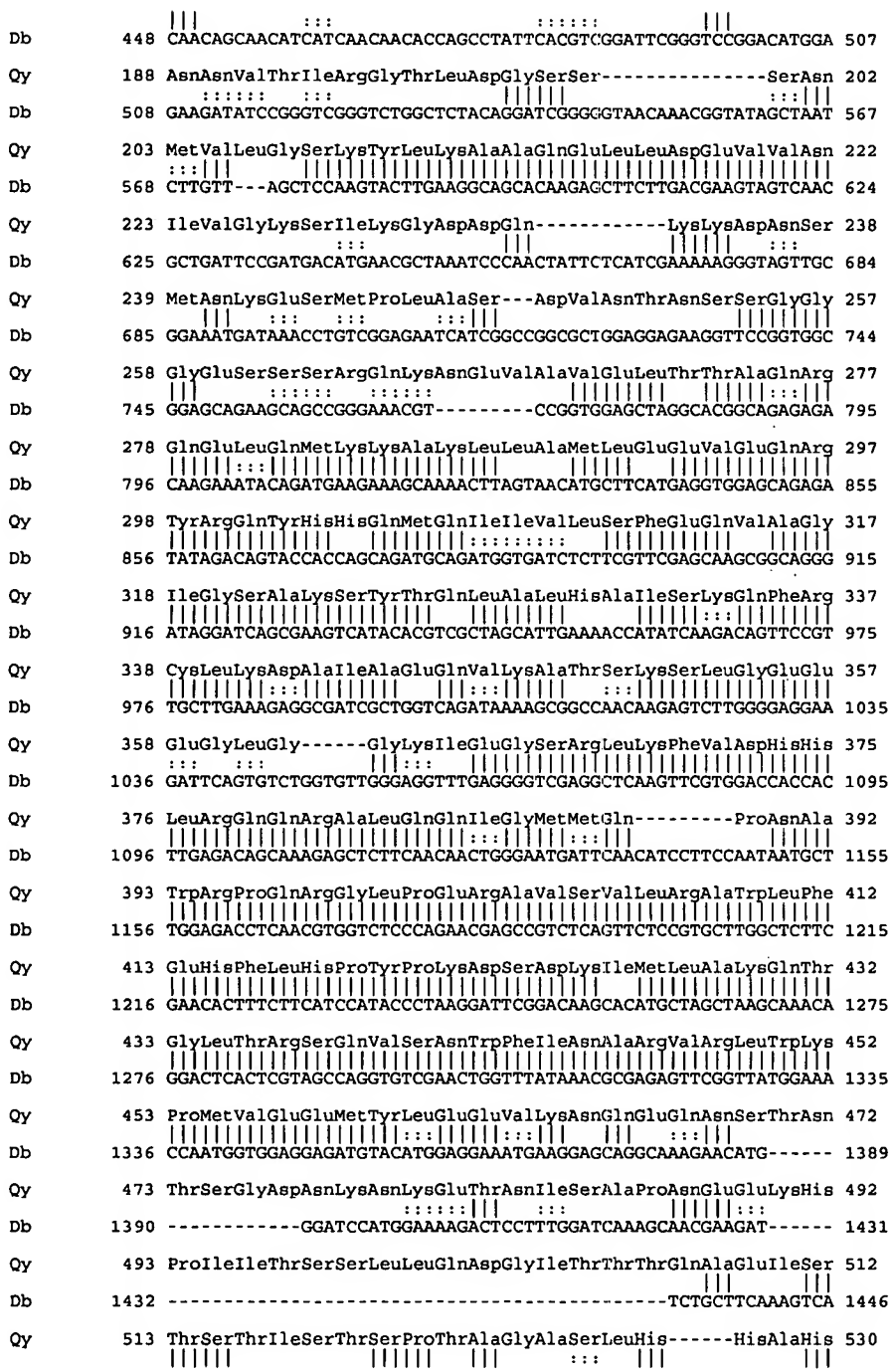
```
Qy        453 ProMetValGluGluMetTyrLeuGluGluValLysAsnGlnGluGlnAsnSerThrAsn 472
              ||||||||||||||||||||:::||||||:::|||   |||   :::|||
Db       1336 CCAATGGTGGAGGAGATGTACATGGAGGAAATGAAGGAGCAGGCAAAGAACATG------ 1389

Qy        473 ThrSerGlyAspAsnLysAsnLysGluThrAsnIleSerAlaProAsnGluGluLysHis 492
                          ::::::|||   :::       ||||||:::
Db       1390 -----------GGATCCATGGAAAAGACTCCTTTGGATCAAAGCAACGAAGAT------ 1431

Qy        493 ProIleIleThrSerSerLeuLeuGlnAspGlyIleThrThrThrGlnAlaGluIleSer 512
                                                          |||      |||
Db       1432 ----------------------------------------TCTGCTTCAAAGTCA 1446

Qy        513 ThrSerThrIleSerThrSerProThrAlaGlyAlaSerLeuHis------HisAlaHis 530
              ||||||          ||||||   |||      :::   |||          |||
Db       1447 ACAAGTAACCAAGAAAAGAGCCCAATGGCGGACACTAATTACCATATGAATCCCAATCAC 1506

Qy        531 AsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrThrValAspHisIle 550
              |||          |||      |||      :::      :::
Db       1507 AACGGTGACCTAGAAGGCGTCACTGGAATGCAAGGAAGCCCCAAGAGACTAAGAACCAGC 1566

Qy        551 GluAsnAsnAlaLysLysGlnArgAsnAspMetHisLysPheSerProSerSerIleLeu 570
              :::          :::      |||          ||||||   :::      |||
Db       1567 GACGAGACAATGATGCAGCCAATAAAT------GCGGATTTCAGCTCCAACGAGAAGCTC 1620

Qy        571 SerSerValAspMetGluAlaLysAlaArgGluSerSerAsnLysGlyPheThrAsnPro 590
              :::        :::|||   :::          |||:::   |||:::      |||
Db       1621 ACGATGAAGATTCTAGAAGAACGGCAAGGGATAAGATCAGATGGTGGCTAC------CCT 1674

Qy        591 LeuMet---------AlaAlaTyrAlaMetGlyAspPheGlyArgPheAsp---ProHis 606
                 |||          |||  |||   ::    ||||||||
Db       1675 TTCATGGGTAATTTCGGGCAATACCAAATGGATGAGATGTCAAGATTTGATGTAGTCTCA 1734

Qy        607 AspGlnGlnMetThrAlaAsnPheHis------GlyAsnAsnGlyValSerLeuThrLeu 624
              ||||||::::::   |||      ::        ||||||||||||||||||||||
Db       1735 GACCAGGAGCTCATGGCGCAAAGGTACTCAGGAAACAACAATGGCGTGTCCCTCACGTTA 1794

Qy        625 GlyLeuProProSerGluAsnLeuAlaMetProValSerGln---------GlnAsnTyr 641
              |||||||||      ::::::|||:::          |||         |||   :::
Db       1795 GGTTTACCTCATTGTGATAGCTTGTCGTCCACGCACCATCAGGGTTTCATGCAGACCCAC 1854

Qy        642 LeuSerAsnAspLeuGlySerArgSerGluMetGly---------------------- 653
                       :::|||   |||   ::::::|||
Db       1855 CATGGGATTCCTATAGGGAGAAGAGTGAAAATAGGAGAAACAGAGGAATATGGACCCGCC 1914

Qy        654 ----------------------------------------------------SerHis 655
                                                                   :::
Db       1915 ACCATCAATGGTGGTAGCTCGACCACAACCGCACATTCATCAGCGGCAGCTGCCGCGGCT 1974

Qy        656 TyrAsnArgMetGlyTyrGluAsnIleAspPheGlnSerGlyAsnLysArgPheProThr 675
              ||||||   |||      :::|||              ||||||:::
Db       1975 TACAATGGGATGAACATACAGAAC-----------------CAGAAGAGATATGTGGCT 2016

Qy        676 GlnLeuLeuProAspPheVal 682
              |||||||||||||||||||
Db       2017 CAGTTATTGCCCGACTTCGTT 2037
```

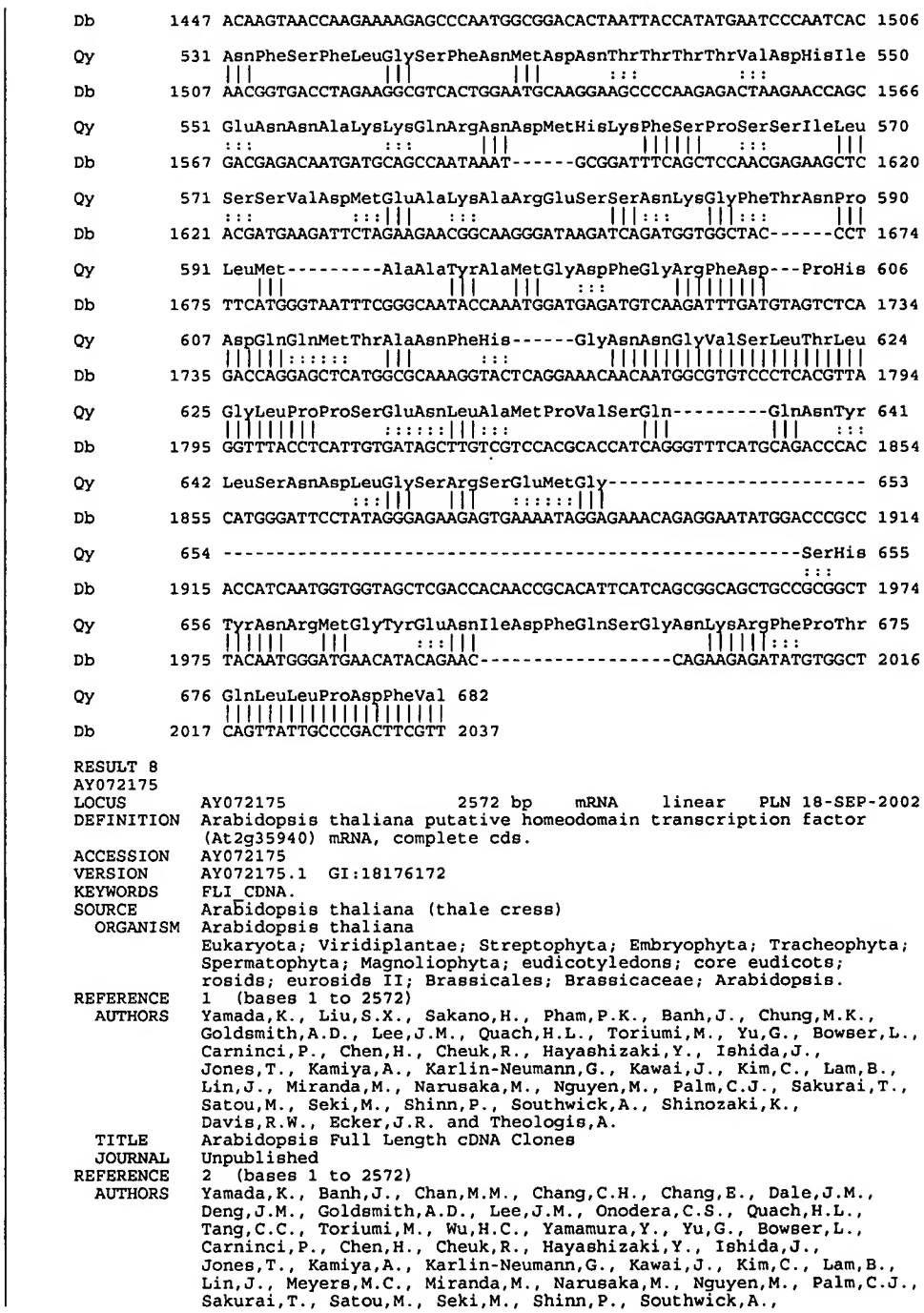
```
RESULT 6
BT008422
LOCUS       BT008422               2043 bp    mRNA    linear   PLN 15-MAY-2003
DEFINITION  Arabidopsis thaliana At2g35940 gene, complete cds.
ACCESSION   BT008422
VERSION     BT008422.1  GI:30725517
KEYWORDS    FLI_CDNA.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 2043)
  AUTHORS   Cheuk,R., Chen,H., Kim,C.J., Shinn,P., Bowser,L., Carninci,P.,
            Dale,J.M., Hayashizaki,Y., Hsuan,V.W., Ishida,J., Jones,T.,
            Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lin,J., Miranda,M.,
            Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L.,
            Sakurai,T., Satou,M., Seki,M., Southwick,A., Toriumi,M., Wong,C.,
            Wu,H.C., Yamada,K., Yu,G., Yuan,S., Shinozaki,K., Davis,R.W.,
            Theologis,A. and Ecker,J.R.
  TITLE     Arabidopsis ORF clones
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2043)
  AUTHORS   Cheuk,R., Chen,H., Kim,C.J., Shinn,P., Bowser,L., Carninci,P.,
            Dale,J.M., Hayashizaki,Y., Hsuan,V.W., Ishida,J., Jones,T.,
            Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lin,J., Miranda,M.,
            Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L.,
            Sakurai,T., Satou,M., Seki,M., Southwick,A., Toriumi,M., Wong,C.,
            Wu,H.C., Yamada,K., Yu,G., Yuan,S., Shinozaki,K., Davis,R.W.,
            Theologis,A. and Ecker,J.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-MAY-2003) Salk Institute Genomic Analysis Laboratory
            (SIGnAL), Plant Biology Laboratory, The Salk Institute for
            Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
            USA
COMMENT     RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
            Arabidopsis Full-Length cDNA') :  Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P.,  Kawai,J.,
            Hayashizaki,Y. and Shinozaki,K.

            The Salk, Stanford, PGEC (SSP) Consortium members constructed and
            sequenced the pUNI (ORF) clones using the RAFL cDNAs: Cheuk,R.,
            Chen,H., Kim,C.J., Shinn,P., Bowser,L., Chan, M.M., Chang,C.M.,
            Dale,J.M., Hsuan,V.W., Jones,T., Karlin-Neumann,G., Lam,B.,
            Lee,J.M., Lin,J., Miranda,M., Nguyen,M., Onodera, C.S., Palm,C.J.,
            Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C.,
            Yamada,K., Yu,G., Yuan,S., Davis,R.W.,  Theologis,A., and
            Ecker,J.R.

            Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to
            this work.  Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)
            contributed equally to this work as PIs.
FEATURES             Location/Qualifiers
     source          1. .2043
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /chromosome="2"
                     /clone="U25364"
                     /ecotype="Columbia"
                     /note="This clone is in pUNI 51"
     CDS             1. .2043
                     /note="putative homeodomain transcription factor"
                     /codon_start=1
                     /product="At2g35940"
                     /protein_id="AAP37781.1"
                     /db_xref="GI:30725518"
                     /translation="MAAYFHGNPPEISAGSDGGLQTLILMNPTTYVQYTQQDNDSNNN
                     NNSNNSNNNNTNTNTNNNNSSFVFLDSHAPQPNASQQFVGIPLSGHEAASITAADNIS
                     VLHGYPPRVQYSLYGSHQVDPTHQQAACETPRAQQGLSLTLSSQQQQQQQHHQQHQPI
                     HVGFGSGHGEDIRVGSGSTGSGVTNGIANLVSSKYLKAAQELLDEVVNADSDDMNAKS
                     QLFSSKKGSCGNDKPVGESSAGAGGEGSGGGAEAAGKRPVELGTAERQEIQMKKAKLS
                     NMLHEVEQRYRQYHQQMQMVISSFEQAAGIGSAKSYTSLALKTISRQFRCLKEAIAGQ
                     IKAANKSLGEEDSVSGVGRFEGSRLKFVDHHLRQQRALQQLGMIQHPSNNAWRPQRGL
                     PERAVSVLRAWLFEHFLHPYPKDSDKHMLAKQTGLTRSQVSNWFINARVRLWKPMVEE
                     MYMEEMKEQAKNMGSMEKTPLDQSNEDSASKSTSNQEKSPMADTNYHMNPNHNGDLEG
                     VTGMQGSPKRLRTSDETMMQPINADFSSNEKLTMKILEERQGIRSDGGYPFMGNFGQY
                     QMDEMSRFDVVSDQELMAQRYSGNNNGVSLTLGLPHCDSLSSTHHQGFMQTHHGIPIG
                     RRVKIGETEEYGPATINGGSSTTTAHSSAAAAAAYNGMNIQNQKRYVAQLLPDFVA"
ORIGIN

Alignment Scores:
Pred. No.:                3.11e-74        Length:          2043
Score:                    1150.50         Matches:         306
Percent Similarity:       51.81%          Conservative:    81
Best Local Similarity:    40.96%          Mismatches:      201
Query Match:              32.29%          Indels:          159
DB:                       8               Gaps:            28
```

US-10-624-201A-2 (1-688) x BT008422 (1-2043)



```
Qy        656 TyrAsnArgMetGlyTyrGluAsnIleAspPheGlnSerGlyAsnLysArgProThr 675
              ||||||  |||      :::|||                    ||||||:::
Db       1975 TACAATGGGATGAACATACAGAAC-----------------CAGAAGAGATATGTGGCT 2016

Qy        676 GlnLeuLeuProAspPheVal 682
              ||||||||||||||||||||
Db       2017 CAGTTATTGCCCGACTTCGTT 2037

RESULT 7
AY096548
LOCUS       AY096548               2074 bp    mRNA    linear   PLN 18-SEP-2002
DEFINITION  Arabidopsis thaliana putative homeodomain transcription factor
            (At2g35940) mRNA, complete cds.
ACCESSION   AY096548
VERSION     AY096548.1  GI:20465476
KEYWORDS    FLI_CDNA.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 2074)
  AUTHORS   Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M.,
            Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L.,
            Tang,C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L.,
            Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
            Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
            Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,
            Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A.,
            Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
  TITLE     Arabidopsis Open Reading Frame (ORF) Clones
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2074)
  AUTHORS   Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M.,
            Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L.,
            Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L.,
            Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
            Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
            Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,
            Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A.,
            Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-APR-2002) Plant Gene Expression Center, 800 Buchanan
            Street, Albany, CA 94710, USA
COMMENT     The RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
            Arabidopsis Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai, J.,
            Hayashizaki,Y. and Shinozaki,K.

            The Salk, Stanford, PGEC (SSP) Consortium members constructed and
            sequenced the pUNI (ORF) clones using the RAFL cDNAs: Yamada,K.,
            Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M.,
            Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C.,
            Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Chen,H.,
            Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Lam,B., Lin,J.,
            Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P.,
            Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.

            Yamada,K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally
            to this work.  Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP
            /PGEC) contributed equally to this work as PIs.

            Annotation is based on the January 2002 version of the Arabidopsis
            genome submitted to GenBank.
FEATURES             Location/Qualifiers
     source          1. .2074
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /chromosome="2"
                     /clone="U18537"
                     /ecotype="Columbia"
                     /note="This clone is in pUNI 51."
     gene            1. .2074
                     /gene="At2g35940"
     CDS             1. .2043
                     /gene="At2g35940"
                     /codon_start=1
                     /evidence=experimental
                     /product="putative homeodomain transcription factor"
                     /protein_id="AAM20198.1"
                     /db_xref="GI:20465477"
                     /translation="MAAYFHGNPPEISAGSDGGLQTLILMNPTTYVQYTQQDNDSNNN
                     NNSNNSNNNNTNTNTNNNNSSFVFLDSHAPQPNASQQFVGIPLSGHEAASITAADNIS
                     VLHGYPPRVQYSLYGSHQVDPTHQQAACETPRAQQGLSLTLSSQQQQQQQHHQQHQPI
                     HVGFGSGHGEDIRVGSGSTGSGVTNGIANLVSSKYLKAAQELLDEVVNADSDDMNAKS
                     QLPSSKKGSCGNDKPVGESSAGAGGEGSGGGAEAAGKRPVELGTAERQEIQMKKAKLS
                     NMLHEVEQRYRQYHQQMQMVISSFEQAAGIGSAKSYTSLALKTISRQFRCLKEAIAGQ
                     IKAANKSLGEEDSVSGVGRFEGSRLKFVDHHLRQQRALQQLGMIQHPSNNAWRPQRGL
                     PERAVSVLRAWLFEHFLHPYPKDSDKHMLAKQTGLTRSQVSNWFINARVRLWKPMVEE
                     MYMEEMKEQAKNMGSMEKTPLDQSNEDSASKSTSNQEKSPMADTNYHMNPNHNGDLEG
                     VTGMQGSPKRLRTSDETMMQPINADFSSNEKLTMKILEERQGIRSDGGYPFMGNFGQY
                     QMDEMSRFDVVSDQELMAQRYSGNNNGVSLTLGLPHCDSLSSTHHQGFMQTHHGIPIG
                     RRVKIGETEEYGPATINGGSSTTTAHSSAAAAAAYNGMNIQNQKRYVAQLLPDFVA"
     3'UTR           2044. .2074
                     /gene="At2g35940"
ORIGIN

Alignment Scores:
Pred. No.:              3.17e-74        Length:         2074
Score:                  1150.50         Matches:        306
Percent Similarity:     51.81%          Conservative:   81
Best Local Similarity:  40.96%          Mismatches:     201
Query Match:            32.29%          Indels:         159
DB:                     8               Gaps:           28

US-10-624-201A-2 (1-688) x AY096548 (1-2074)

Qy         22 HisGlyAsn-----------------SerAsnAsnAsnAsnIleGlnThrLeuTyrLeu 35
              |||||||||                 ::::::      :::|||||||||   |||
Db         16 CACGGAAACCCACCGGAGATCTCTGCCGGATCCGACGGTGGTCTTCAAACGTTGATCCTC 75

Qy         36 MetAsnProAsnAsnTyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGlnGlnGln 55
              |||||||||      |||:::|||   ||||||      |||
Db         76 ATGAATCCAACTACTTACGTTCAG---TACACCCAACAAGACAACGACTCGAACAACAAC 132

Qy         56 LeuLeuPheLeuAsnSerSerProAlaAlaSerAsnAlaLeuCysHisAlaAsnIle--- 74
                    |||:::|||      ::::::|||          :::   |||
Db        133 -----------AACAACAGCAACAATAGCAACAACAACAACACAAACACAAACAC 180

Qy         75 ----------------------------GlnHisAlaProLeu------GlnGlnGln 82
                                          |||||||||             |||
Db        181 AACAACAACAGTAGTTTCGTTTTCCTCGATTCCCACGCGCCGCAGCCAAACGCGAGCCAG 240

Qy         83 HisPheValGlyValProLeu---------------ProAlaValSerLeuHisAspGln 97
                 |||||||||:::||||||                    :::::::::      |||
Db        241 CAGTTCGTCGGAATACCACTCTCAGGTCACGAAGCTGCTTCCATTACAGCCGCCGACAAC 300

Qy         98 IleAsn---HisHisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnSerGlnGln 116
              |||:::     ||||||        |||:::     :::         |||   |||
Db        301 ATCTCCGTACTTCACGGTTATCCTCCGCGCGTGCAGTACAGTCTCTACGGTAGCCACCAA 360

Qy        117 ValIleValProSerSerThrGlyValSerAlaThrSerCysGlyGlyIleThrThrAsp 136
              |||
Db        361 GTG---------------------------------------------------------- 363

Qy        137 LeuAlaSerGlnLeuAlaPheGlnArgProIleProThrProGlnHisArgGln------ 154
                                                  |||   |||:::|||
Db        364 ----------------------------------GATCCCACTCACCAGCAAGCCGCG 387

Qy        155 ------GlnGlnGlnGlnGlnGlyGlyLeuSerLeuSerLeuSerProGlnLeuGlnGln 172
                    :::   |||  ||||||||||||:::||||||   |||   ||||||
Db        388 TGTGAGACGCCACGCGCGCAGCAAGGCCTCTCTTTAACCCTCTCGTCTCAACAGCAGCAG 447

Qy        173 GlnIleSerPheAsnAsn---------------AsnIleSerSerSerSerProArgThr 187
```

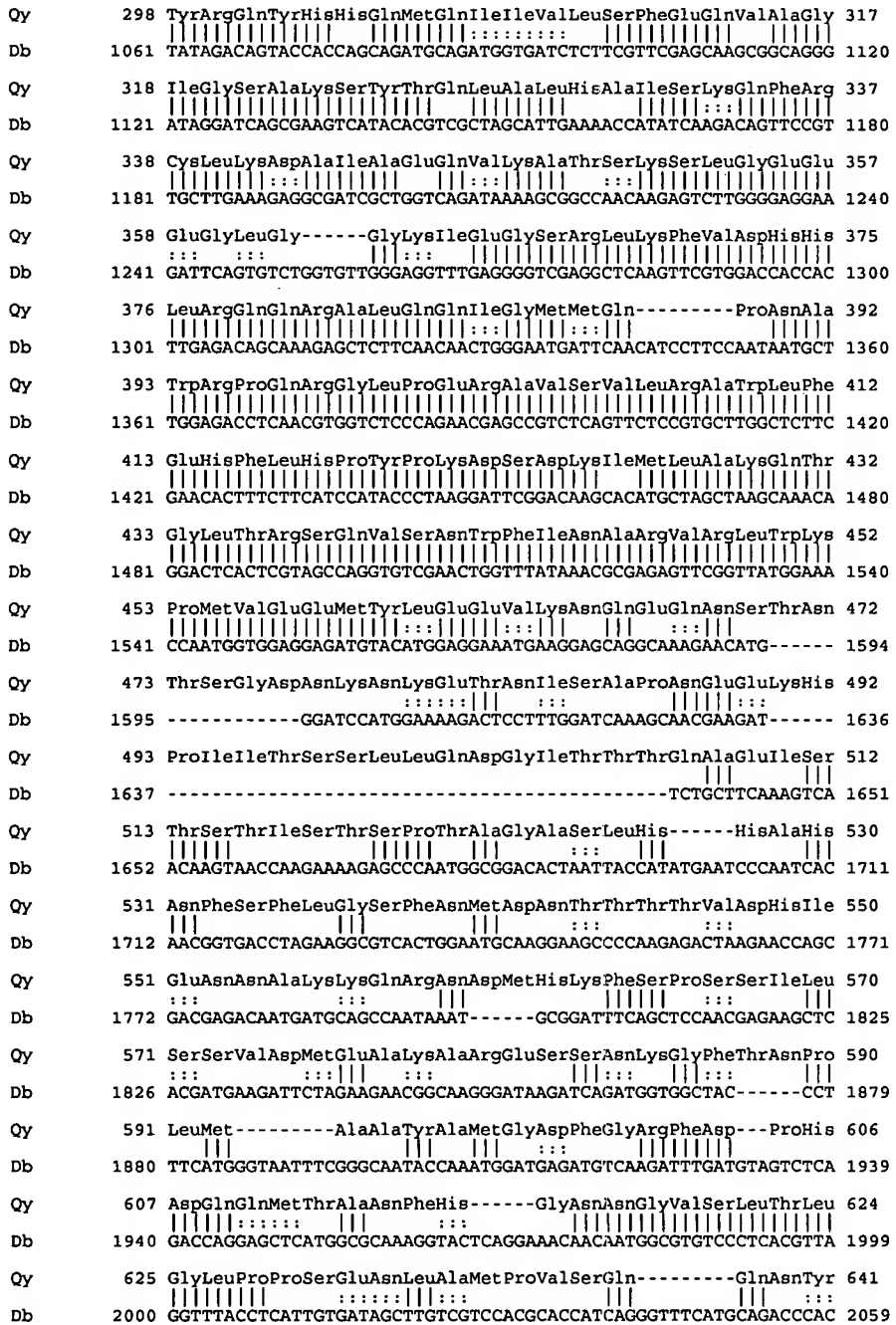




RESULT 8
AY072175
LOCUS AY072175 2572 bp mRNA linear PLN 18-SEP-2002
DEFINITION Arabidopsis thaliana putative homeodomain transcription factor (At2g35940) mRNA, complete cds.
ACCESSION AY072175
VERSION AY072175.1 GI:18176172
KEYWORDS FLI_CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 2572)
AUTHORS Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K., Goldsmith,A.D., Lee,J.M., Quach,H.L., Toriumi,M., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
TITLE Arabidopsis Full Length cDNA Clones
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2572)
AUTHORS Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A.,

```
          Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
  TITLE   Direct Submission
  JOURNAL Submitted (02-JAN-2002) Plant Gene Expression Center, 800 Buchanan
          Street, Albany, CA 94710, USA
COMMENT   RIKEN Genomic Sciences Center (GSC) members carried out the
          collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
          Arabidopsis Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
          Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
          Hayashizaki,Y. and Shinozaki,K.

          The Salk, Stanford, PGEC (SSP) Consortium members carried out the
          sequencing and annotation of the RAFL cDNAs: Yamada,K., Banh,J.,
          Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M.,
          Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C.,
          Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Chen,H.,
          Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Lam,B., Lin,J.,
          Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P.,
          Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.

          Yamada,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally to
          this work.  Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC)
          contributed equally to this work as PIs.

          Annotation is based on the January 2002 version of the Arabidopsis
          genome submitted to GenBank.
FEATURES             Location/Qualifiers
     source          1. .2572
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /chromosome="2"
                     /clone="RAFL09-25-N04 (R18537)"
                     /ecotype="Columbia"
                     /note="This clone is in a modified pBluescript vector
                     (FLC-1) as a BamHI/XhoI insert."
     gene            1. .2572
                     /gene="At2g35940"
     5'UTR           1. .205
                     /gene="At2g35940"
     misc_difference 2
                     /gene="At2g35940"
                     /note="compared to genomic sequence"
                     /replace="c"
     CDS             206. .2248
                     /gene="At2g35940"
                     /codon_start=1
                     /evidence=experimental
                     /product="putative homeodomain transcription factor"
                     /protein_id="AAL59997.1"
                     /db_xref="GI:18176173"
                     /translation="MAAYFHGNPPEISAGSDGGLQTLILMNPTTYVQYTQQDNDSNNN
                     NNSNNSNNNNTNTNTNNNNSSFVFLDSHAPQPNASQQFVGIPLSGHEAASITAADNIS
                     VLHGYPPRVQYSLYGSHQVDPTHQQAACETPRAQQGLSLTLSSQQQQQQQHHQQHQPI
                     HVGFGSGHGEDIRVGSGSTGSGVTNGIANLVSSKYLKAAQELLDEVVNADSDDMNAKS
                     QLFSSKKGSCGNDKPVGESSAGAGGEGSGGGAEAAGKRPVELGTAERQEIQMKKAKLS
                     NMLHEVEQRYRQYHQQMQMVISSFEQAAGIGSAKSYTSLALKTISRQFRCLKEAIAGQ
                     IKAANKSLGEEDSVSGVGRFEGSRLKFVDHHLRQQRALQQLGMIQHPSNNAWRPQRGL
                     PERAVSVLRAWLFEHFLHPYPKDSDKHMLAKQTGLTRSQVSNWFINARVRLWKPMVEE
                     MYMEEMKEQAKNMGSMEKTPLDQSNEDSASKSTSNQEKSPMADTNYHMNPNHNGDLEG
                     VTGMQGSPKRLRTSDETMMQPINADFSSNEKLTMKILEERQGIRSDGGYPFMGNFGQY
                     QMDEMSRFDVVSDQELMAQRYSGNNNGVSLTLGLPHCDSLSSTHHQGFMQTHHGIPIG
                     RRVKIGETEEYGPATINGGSSTTTAHSSAAAAAAYNGMNIQNQKRYVAQLLPDFVA"
     3'UTR           2249. .2572
                     /gene="At2g35940"
     misc_difference 2417. .2418
                     /gene="At2g35940"
                     /note="compared to genomic sequence"
                     /replace="ttc"
     misc_difference 2556
                     /gene="At2g35940"
                     /note="not present in genomic sequence"
ORIGIN
```

```
Alignment Scores:
Pred. No.:              4.19e-74        Length:         2572
Score:                  1150.50         Matches:        306
Percent Similarity:     51.81%          Conservative:   81
Best Local Similarity:  40.96%          Mismatches:     201
Query Match:            32.29%          Indels:         159
DB:                     8               Gaps:           28

US-10-624-201A-2 (1-688) x AY072175 (1-2572)

Qy         22 HisGlyAsn----------------SerAsnAsnAsnAsnIleGlnThrLeuTyrLeu 35
              |||||||||                ::::::     :::|||||||||   |||
Db        221 CACGGAAACCCACCGGAGATCTCTGCCGGATCCGACGGTGGTCTTCAAACGTTGATCCTC 280

Qy         36 MetAsnProAsnAsnTyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGlnGlnGln 55
              |||||||||      |||:::|||   ||||||      |||
Db        281 ATGAATCCAACTACTTACGTTCAG---TACACCCAACAAGACAACGACTCGAACAACAAC 337

Qy         56 LeuLeuPheLeuAsnSerSerProAlaAlaSerAsnAlaLeuCysHisAlaAsnIle--- 74
                        |||:::|||      ::::::|||          :::   |||
Db        338 -----------AACAACAGCAACAATAGCAACAACAACAACACAAACACAAACACAAAC 385

Qy         75 ----------------------------GlnHisAlaProLeu------GlnGlnGln 82
                                          |||||||||              |||
Db        386 AACAACAACAGTAGTTTCGTTTTCCTCGATTCCCACGCGCCGCAGCCAAACGCGAGCCAG 445

Qy         83 HisPheValGlyValProLeu--------------ProAlaValSerLeuHisAspGln 97
                 |||||||||:::||||||                 :::::::::      |||
Db        446 CAGTTCGTCGGAATACCACTCTCAGGTCACGAAGCTGCTTCCATTACAGCCGCCGACAAC 505

Qy         98 IleAsn---HisHisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnSerGlnGln 116
              |||:::      ||||||         |||:::      :::         |||   |||
Db        506 ATCTCCGTACTTCACGGTTATCCTCCGCGCGTGCAGTACAGTCTCTACGGTAGCCACCAA 565

Qy        117 ValIleValProSerSerThrGlyValSerAlaThrSerCysGlyGlyIleThrThrAsp 136
              |||
Db        566 GTG---------------------------------------------------------- 568

Qy        137 LeuAlaSerGlnLeuAlaPheGlnArgProIleProThrProGlnHisArgGln------ 154
                                                  |||   |||:::|||
Db        569 ----------------------------------GATCCCACTCACCAGCAAGCCGCG 592

Qy        155 ------GlnGlnGlnGlnGlnGlyGlyLeuSerLeuSerLeuSerProGlnLeuGlnGln 172
                    :::   |||   ||||||||||||:::||||||   |||   ||||||
Db        593 TGTGAGACGCCACGCGCGCAGCAAGGCCTCTCTTTAACCCTCTCGTCTCAACAGCAGCAG 652

Qy        173 GlnIleSerPheAsnAsn----------------AsnIleSerSerSerSerProArgThr 187
              |||         :::                  ::::::          |||
Db        653 CAACAGCAACATCATCAACAACACCAGCCTATTCACGTCGGATTCGGGTCCGGACATGGA 712

Qy        188 AsnAsnValThrIleArgGlyThrLeuAspGlySerSer---------------SerAsn 202
                 ::::::   :::                 ||||||                ::|||
Db        713 GAAGATATCCGGGTCGGGTCTGGCTCTACAGGATCGGGGGTAACAAACGGTATAGCTAAT 772

Qy        203 MetValLeuGlySerLysTyrLeuLysAlaAlaGlnGluLeuLeuAspGluValValAsn 222
              :::|||      ||||||||||||||||||||||||||||||||||||||||||||||
Db        773 CTTGTT---AGCTCCAAGTACTTGAAGGCAGCACAAGAGCTTCTTGACGAAGTAGTCAAC 829

Qy        223 IleValGlyLysSerIleLysGlyAspAspGln------------LysLysAspAsnSer 238
                          :::              |||            ||||||   :::
Db        830 GCTGATTCCGATGACATGAACGCTAAATCCCAACTATTCTCATCGAAAAAGGGTAGTTGC 889

Qy        239 MetAsnLysGluSerMetProLeuAlaSer---AspValAsnThrAsnSerSerGlyGly 257
                 |||   :::   :::      :::|||                        |||||||||
Db        890 GGAAATGATAAACCTGTCGGAGAATCATCGGCCGGCGCTGGAGGAGAAGGTTCCGGTGGC 949

Qy        258 GlyGluSerSerSerArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGlnArg 277
              |||      ::::::   ::::::          |||||||||   ||||||:::|||
Db        950 GGAGCAGAAGCAGCCGGGAAACGT---------CCGGTGGAGCTAGGCACGGCAGAGAGA 1000

Qy        278 GlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeuGluGluValGluGlnArg 297
              ||||||:::|||||||||||||||||||||      ||||||   ||||||||||||||||
Db       1001 CAAGAAATACAGATGAAGAAAGCAAAACTTAGTAACATGCTTCATGAGGTGGAGCAGAGA 1060
```



RESULT 9
AY099854
LOCUS AY099854 2578 bp mRNA linear PLN 06-MAY-2002
DEFINITION Arabidopsis thaliana putative homeodomain transcription factor (At2g35940) mRNA, complete cds.
ACCESSION AY099854
VERSION AY099854.1 GI:20466775
KEYWORDS FLI_CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 2578)
AUTHORS Nguyen,M., Karlin-Neumann,G., Southwick,A., Lam,B., Miranda,M., Palm,C.J., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H., Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A., Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K., Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K., Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.
TITLE Direct Submission
JOURNAL Submitted (24-APR-2002) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
COMMENT e-mail for correspondence: arab@sequence.stanford.edu

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Nguyen,M., Southwick,A., Karlin-Neumann,G., Lam,B., Miranda,M., Palm,C.J., Bowser,L., Jones,T., Banh,J., Chen,H., Cheuk,R., Chung,M.K., Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H., Shinn,P., Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.

Nguyen,M, (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W. (SSP/Stanford) contributed equally to this work as PIs.

FEATURES Location/Qualifiers
source 1. .2578
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="2"
/clone="RAFL09-94-K12 (R25364)"
/ecotype="Columbia"
/note="This clone is in pBluescript"
gene 1. .2578
/gene="At2g35940"
/note="synonym: F11F19.15"
CDS 198. .2240
/gene="At2g35940"
/codon_start=1

```
                     /product="putative homeodomain transcription factor"
                     /protein_id="AAM20705.1"
                     /db_xref="GI:20466776"
                     /translation="MAAYFHGNPPEISAGSDGGLQTLILMNPTTYVQYTQQDNDSNNN
                     NNSNNSNNNNTNTNTNNNNSSFVFLDSHAPQPNASQQFVGIPLSGHEAASITAADNIS
                     VLHGYPPRVQYSLYGSHQVDPTHQQAACETPRAQQGLSLTLSSQQQQQQQHHQQHQPI
                     HVGFGSGHGEDIRVGSGSTGSGVTNGIANLVSSKYLKAAQELLDEVVNADSDDMNAKS
                     QLFSSKKGSCGNDKPVGESSAGAGGEGSGGGAEAAGKRPVELGTAERQEIQMKKAKLS
                     NMLHEVEQRYRQYHQQMQMVISSFEQAAGIGSAKSYTSLALKTISRQFRCLKEAIAGQ
                     IKAANKSLGEEDSVSGVGRFEGSRLKFVDHHLRQQRALQQLGMIQHPSNNAWRPQRGL
                     PERAVSVLRAWLFEHFLHPYPKDSDKHMLAKQTGLTRSQVSNWFINARVRLWKPMVEE
                     MYMEEMKEQAKNMGSMEKTPLDQSNEDSASKSTSNQEKSPMADTNYHMNPNHNGDLEG
                     VTGMQGSPKRLRTSDETMMQPINADFSSNEKLTMKILEERQGIRSDGGYPFMGNFGQY
                     QMDEMSRFDVVSDQELMAQRYSGNNNGVSLTLGLPHCDSLSSTHHQGFMQTHHGIPIG
                     RRVKIGETEEYGPATINGGSSTTTAHSSAAAAAAYNGMNIQNQKRYVAQLLPDFVA"
ORIGIN

Alignment Scores:
Pred. No.:              4.21e-74      Length:         2578
Score:                  1150.50       Matches:        306
Percent Similarity:     51.81%        Conservative:   81
Best Local Similarity:  40.96%        Mismatches:     201
Query Match:            32.29%        Indels:         159
DB:                     8             Gaps:           28

US-10-624-201A-2 (1-688) x AY099854 (1-2578)

Qy         22 HisGlyAsn-----------------SerAsnAsnAsnAsnIleGlnThrLeuTyrLeu 35
              |||||||||                      ::::::     :::|||||||||   |||
Db        213 CACGGAAACCCACCGGAGATCTCTGCCGGATCCGACGGTGGTCTTCAAACGTTGATCCTC 272

Qy         36 MetAsnProAsnAsnTyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGlnGlnGln 55
              |||||||||      |||:::|||   ||||||      |||
Db        273 ATGAATCCAACTACTTACGTTCAG---TACACCCAACAAGACAACGACTCGAACAACAAC 329

Qy         56 LeuLeuPheLeuAsnSerSerProAlaAlaSerAsnAlaLeuCysHisAlaAsnIle--- 74
                        |||:::|||      ::::::|||         :::   |||
Db        330 -----------AACAACAGCAACAATAGCAACAACAACAACACAAACACAAACACAAAC 377

Qy         75 ------------------------------GlnHisAlaProLeu------GlnGlnGln 82
                                            |||||||||               |||
Db        378 AACAACAACAGTAGTTTCGTTTTCCTCGATTCCCACGCGCCGCAGCCAAACGCGAGCCAG 437

Qy         83 HisPheValGlyValProLeu---------------ProAlaValSerLeuHisAspGln 97
                 |||||||||:::||||||                    :::::::::      |||
Db        438 CAGTTCGTCGGAATACCACTCTCAGGTCACGAAGCTGCTTCCATTACAGCCGCCGACAAC 497

Qy         98 IleAsn---HisHisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnSerGlnGln 116
              |||:::      ||||||         |||:::      :::         |||   |||
Db        498 ATCTCCGTACTTCACGGTTATCCTCCGCGCGTGCAGTACAGTCTCTACGGTAGCCACCAA 557

Qy        117 ValIleValProSerSerThrGlyValSerAlaThrSerCysGlyGlyIleThrThrAsp 136
              |||
Db        558 GTG------------------------------------------------------- 560

Qy        137 LeuAlaSerGlnLeuAlaPheGlnArgProIleProThrProGlnHisArgGln------ 154
                                                 |||   |||:::|||
Db        561 -----------------------------------GATCCCACTCACCAGCAAGCCGCG 584

Qy        155 ------GlnGlnGlnGlnGlnGlyGlyLeuSerLeuSerLeuSerProGlnLeuGlnGln 172
                    :::   |||   |||||||||||||:::||||||   |||   ||||||
Db        585 TGTGAGACGCCACGCGCGCAGCAAGGCCTCTCTTTAACCCTCTCGTCTCAACAGCAGCAG 644

Qy        173 GlnIleSerPheAsnAsn---------------AsnIleSerSerSerSerProArgThr 187
              |||         :::                 ::::::          |||
Db        645 CAACAGCAACATCATCAACAACACCAGCCTATTCACGTCGGATTCGGGTCCGGACATGGA 704

Qy        188 AsnAsnValThrIleArgGlyThrLeuAspGlySerSer---------------SerAsn 202
                 ::::::   :::                |||||                  :::|||
Db        705 GAAGATATCCGGGTCGGGTCTGGCTCTACAGGATCGGGGGTAACAAACGGTATAGCTAAT 764

Qy        203 MetValLeuGlySerLysTyrLeuLysAlaAlaGlnGluLeuLeuAspGluValValAsn 222
              :::|||      ||||||||||||||||||||||||||||||||||||||||||||||||

Db        765 CTTGTT---AGCTCCAAGTACTTGAAGGCAGCACAAGAGCTTCTTGACGAAGTAGTCAAC 821

Qy        223 IleValGlyLysSerIleLysGlyAspAspGln-----------LysLysAspAsnSer 238
                          :::             |||           ||||||   :::
Db        822 GCTGATTCCGATGACATGAACGCTAAATCCCAACTATTCTCATCGAAAAAGGGTAGTTGC 881

Qy        239 MetAsnLysGluSerMetProLeuAlaSer---AspValAsnThrAsnSerSerGlyGly 257
                 |||   :::   :::      :::|||                     |||||||||
Db        882 GGAAATGATAAACCTGTCGGAGAATCATCGGCCGGCGCTGGAGGAGAAGGTTCCGGTGGC 941

Qy        258 GlyGluSerSerSerArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGlnArg 277
              |||      ::::::  ::::::           |||||||||   ||||||:::|||
Db        942 GGAGCAGAAGCAGCCGGGAAACGT---------CCGGTGGAGCTAGGCACGGCAGAGAGA 992

Qy        278 GlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeuGluGluValGluGlnArg 297
              ||||||:::|||||||||||||||||||||      ||||||   ||||||||||||||||
Db        993 CAAGAAATACAGATGAAGAAAGCAAAACTTAGTAACATGCTTCATGAGGTGGAGCAGAGA 1052

Qy        298 TyrArgGlnTyrHisHisGlnMetGlnIleIleValLeuSerPheGluGlnValAlaGly 317
              ||||||||||||||   |||||||||:::::::::   ||||||||||   ||||||
Db       1053 TATAGACAGTACCACCAGCAGATGCAGATGGTGATCTCTTCGTTCGAGCAAGCGGCAGGG 1112

Qy        318 IleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArg 337
              ||||||||||||||||||||||    |||||||||      ||||||:::|||||||||
Db       1113 ATAGGATCAGCGAAGTCATACACGTCGCTAGCATTGAAAACCATATCAAGACAGTTCCGT 1172

Qy        338 CysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGluGlu 357
              |||||||||:::|||||||||   |||:::||||||   :::|||||||||||||||||||
Db       1173 TGCTTGAAAGAGGCGATCGCTGGTCAGATAAAAGCGGCCAACAAGAGTCTTGGGGAGGAA 1232

Qy        358 GluGlyLeuGly------GlyLysIleGluGlySerArgLeuLysPheValAspHisHis 375
              :::  :::          |||:::  ||||||||||||||||||||||||||||||||||
Db       1233 GATTCAGTGTCTGGTGTTGGGAGGTTTGAGGGGTCGAGGCTCAAGTTCGTGGACCACCAC 1292

Qy        376 LeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGln---------ProAsnAla 392
              |||||||||||||||||||||||||:::||||||:::|||            ||||||
Db       1293 TTGAGACAGCAAAGAGCTCTTCAACAACTGGGAATGATTCAACATCCTTCCAATAATGCT 1352

Qy        393 TrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeuPhe 412
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1353 TGGAGACCTCAACGTGGTCTCCCAGAACGAGCCGTCTCAGTTCTCCGTGCTTGGCTCTTC 1412

Qy        413 GluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGlnThr 432
              ||||||||||||||||||||||||||||||||||||   ||||||||||||||||||||
Db       1413 GAACACTTTCTTCATCCATACCCTAAGGATTCGGACAAGCACATGCTAGCTAAGCAAACA 1472

Qy        433 GlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLys 452
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1473 GGACTCACTCGTAGCCAGGTGTCGAACTGGTTTATAAACGCGAGAGTTCGGTTATGGAAA 1532

Qy        453 ProMetValGluGluMetTyrLeuGluGluValLysAsnGlnGluGlnAsnSerThrAsn 472
              ||||||||||||||||||:::||||||:::|||   |||   :::|||
Db       1533 CCAATGGTGGAGGAGATGTACATGGAGGAAATGAAGGAGCAGGCAAAGAACATG------ 1586

Qy        473 ThrSerGlyAspAsnLysAsnLysGluThrAsnIleSerAlaProAsnGluGluLysHis 492
                                  ::::::|||   :::         ||||||:::
Db       1587 -----------GGATCCATGGAAAAGACTCCTTTGGATCAAAGCAACGAAGAT------ 1628

Qy        493 ProIleIleThrSerSerLeuGlnAspGlyIleThrThrThrGlnAlaGluIleSer 512
                                                                |||   |||
Db       1629 ------------------------------------------TCTGCTTCAAAGTCA 1643

Qy        513 ThrSerThrIleSerThrSerProThrAlaGlyAlaSerLeuHis------HisAlaHis 530
              ||||||           ||||||  |||      :::  |||           |||
Db       1644 ACAAGTAACCAAGAAAAGAGCCCAATGGCGGACACTAATTACCATATGAATCCCAATCAC 1703

Qy        531 AsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrThrValAspHisIle 550
              |||          |||      |||      :::      :::
Db       1704 AACGGTGACCTAGAAGGCGTCACTGGAATGCAAGGAAGCCCCAAGAGACTAAGAACCAGC 1763

Qy        551 GluAsnAsnAlaLysLysGlnArgAsnAspMetHisLysPheSerProSerSerIleLeu 570
              :::          :::      |||          ||||||  :::      |||
Db       1764 GACGAGACAATGATGCAGCCAATAAAT------GCGGATTTCAGCTCCAACGAGAAGCTC 1817
```

```
Qy        571 SerSerValAspMetGluAlaLysAlaArgGluSerSerAsnLysGlyPheThrAsnPro 590
Db       1818 ACGATGAAGATTCTAGAAGAACGGCAAGGGATAAGATCAGATGGTGGCTAC------CCT 1871

Qy        591 LeuMet---------AlaAlaTyrAlaMetGlyAspPheGlyArgPheAsp---ProHis 606
Db       1872 TTCATGGGTAATTTCGGGCAATACCAAATGGATGAGATGTCAAGATTTGATGTAGTCTCA 1931

Qy        607 AspGlnGlnMetThrAlaAsnPheHis------GlyAsnAsnGlyValSerLeuThrLeu 624
Db       1932 GACCAGGAGCTCATGGCGCAAAGGTACTCAGGAAACAACAATGGCGTGTCCCTCACGTTA 1991

Qy        625 GlyLeuProProSerGluAsnLeuAlaMetProValSerGln---------GlnAsnTyr 641
Db       1992 GGTTTACCTCATTGTGATAGCTTGTCGTCCACGCACCATCAGGGTTTCATGCAGACCCAC 2051

Qy        642 LeuSerAsnAspLeuGlySerArgSerGluMetGly---------------------- 653
Db       2052 CATGGGATTCCTATAGGGAGAAGAGTGAAAATAGGAGAAACAGAGGAATATGGACCCGCC 2111

Qy        654 ----------------------------------------------------SerHis 655
Db       2112 ACCATCAATGGTGGTAGCTCGACCACAACCGCACATTCATCAGCGGCAGCTGCCGCGGCT 2171

Qy        656 TyrAsnArgMetGlyTyrGluAsnIleAspPheGlnSerGlyAsnLysArgProThr 675
Db       2172 TACAATGGGATGAACATACAGAAC-----------------CAGAAGAGATATGTGGCT 2213

Qy        676 GlnLeuLeuProAspPheVal 682
Db       2214 CAGTTATTGCCCGACTTCGTT 2234

RESULT 10
AF353094
LOCUS       AF353094               2385 bp    mRNA    linear   PLN 30-APR-2001
DEFINITION  Arabidopsis thaliana BEL1-like homeodomain 1 (BLH1) mRNA, complete
            cds.
ACCESSION   AF353094
VERSION     AF353094.1  GI:13877516
KEYWORDS    .
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 2385)
  AUTHORS   Pidkowich,M.S., Samach,A., Modrusan,Z. and Haughn,G.W.
  TITLE     A family of BEL1-like homeodomain (BLH) proteins in Arabidopsis
            thaliana
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2385)
  AUTHORS   Pidkowich,M.S., Samach,A., Modrusan,Z. and Haughn,G.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-FEB-2001) Botany, University of British Columbia,
            #3529 - 6270 University Blvd., Vancouver, BC V6T 1Z4, Canada
FEATURES             Location/Qualifiers
     source          1. .2385
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /chromosome="II"
     gene            1. .2385
                     /gene="BLH1"
     CDS             179. .2221
                     /gene="BLH1"
                     /note="homeodomain protein BLH1; transcription factor;
                     capable of heterodimerization with KNAT proteins; similar
                     to predicted homeodomain protein presented in GenBank
                     Accession Number AAD21463"
                     /codon_start=1
                     /product="BEL1-like homeodomain 1"
                     /protein_id="AAK43836.1"
                     /db_xref="GI:13877517"
                     /translation="MAAYFHGNPPEISAGSDGGLQTLILMNPTTYVQYTQQDNDSNNN
                     NNSNNSNNNNTNTNTNNNNSSFVFLDSHAPQPNASQQFVGIPLSGHEAASITAADNIS
                     VLHGYPPRVQYSLYGSHQVDPTHQQAACETPRAQQGLSLTLSSQQQQQQQHHQQHQPI
                     HVGFGSGHGEDIRVGSGSTGSGVTNGIANLVSSKYLKAAQELLDEVVNADSDDMNAKS
                     QLFSSKKGSCGNDKPVGESSAGAGGEGSGGGAEAAGKRPVELGTAERQEIQMKKAKLS
                     NMLHEVEQRYRQYHQQMQMVISSFEQAAGIGSAKSYTSLALKTISRQFRCLKEAIAGQ
                     IKAANKSLGEEDSVSGVGRFEGSRLKFVDHHLRQQRALQQLGMIQHPSNNAWRPQRGL
                     PERAVSVLRAWLFEHFLHPYPKDSDKHMLAKQTGLTRSQVSNWFINARVRLWKPMVEE
                     MYMEEMKEQAKNMGSMEKTPLDQSNEDSASKSTSNQEKSPMADTNYHMNPNHNGDLEG
                     VTGMQGCPKRLRTSDETMMQPINADFSSNEKLTMKILEERQGIRSDGGYPFMGNFGQY
                     QMDEMSRFDVVSDQELMAQRYSGNNNGVSLTLGLPHCDSLSSTDHQGFMQTHHGIPIG
                     RRVKIGETEEYGPATINGGSSTTTAHSSAAAAAAYNGMNIQNQKRYVAQLLPDFVA"
ORIGIN

Alignment Scores:
Pred. No.:                5.31e-74        Length:          2385
Score:                    1148.50         Matches:         306
Percent Similarity:       51.67%          Conservative:    80
Best Local Similarity:    40.96%          Mismatches:      202
Query Match:              32.23%          Indels:          159
DB:                       8               Gaps:            28

US-10-624-201A-2 (1-688) x AF353094 (1-2385)

Qy         22 HisGlyAsn-----------------SerAsnAsnAsnAsnIleGlnThrLeuTyrLeu 35
Db        194 CACGGAAACCCACCGGAGATCTCTGCCGGATCCGACGGTGGTCTTCAAACGTTGATCCTC 253

Qy         36 MetAsnProAsnAsnTyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGlnGlnGln 55
Db        254 ATGAATCCAACTACTTACGTTCAG---TACACCCAACAAGACAACGACTCGAACAACAAC 310

Qy         56 LeuLeuPheLeuAsnSerSerProAlaAlaSerAsnAlaLeuCysHisAlaAsnIle--- 74
Db        311 -----------AACAACAGCAACAATAGCAACAACAACAACACAAACACAAACACAAAC 358

Qy         75 ----------------------------GlnHisAlaProLeu------GlnGlnGln 82
Db        359 AACAACAACAGTAGTTTCGTTTTCCTCGATTCCCACGCGCCGCAGCCAAACGCGAGCCAG 418

Qy         83 HisPheValGlyValProLeu--------------ProAlaValSerLeuHisAspGln 97
Db        419 CAGTTCGTCGGAATACCACTCTCAGGTCACGAAGCTGCTTCCATTACAGCCGCCGACAAC 478

Qy         98 IleAsn---HisHisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAsnAspGlnSerGlnGln 116
Db        479 ATCTCCGTACTTCACGGTTATCCTCCGCGCGTGCAGTACAGTCTCTACGGTAGCCACCAA 538

Qy        117 ValIleValProSerSerThrGlyValSerAlaThrSerCysGlyGlyIleThrThrAsp 136
Db        539 GTG---------------------------------------------------------- 541

Qy        137 LeuAlaSerGlnLeuAlaPheGlnArgProIleProThrProGlnHisArgGln------ 154
Db        542 ----------------------------------GATCCCACTCACCAGCAAGCCGCG 565

Qy        155 ------GlnGlnGlnGlnGlnGlyGlyLeuSerLeuSerLeuSerProGlnLeuGlnGln 172
Db        566 TGTGAGACGCCACGCGCGCAGCAAGGCCTCTCTTTAACCCTCTCGTCTCAACAGCAGCAG 625

Qy        173 GlnIleSerPheAsnAsn--------------AsnIleSerSerSerSerProArgThr 187
Db        626 CAACAGCAACATCATCAACAACACCAGCCTATTCACGTCGGATTCGGGTCCGGACATGGA 685

Qy        188 AsnAsnValThrIleArgGlyThrLeuAspGlySerSer--------------SerAsn 202
Db        686 GAAGATATCCGGGTCGGGTCTGGCTCTACAGGATCGGGGGTAACAAACGGTATAGCTAAT 745

Qy        203 MetValLeuGlySerLysTyrLeuLysAlaAlaGlnGluLeuLeuAspGluValValAsn 222
Db        746 CTTGTT---AGCTCCAAGTACTTGAAGGCAGCACAAGAGCTTCTTGACGAAGTAGTCAAC 802
```

```
Qy       223 IleValGlyLysSerIleLysGlyAspAspGln-----------LysLysAspAsnSer 238
Db       803 GCTGATTCCGATGACATGAACGCTAAATCCCAACTATTCTCATCGAAAAAGGGTAGTTGC 862

Qy       239 MetAsnLysGluSerMetProLeuAlaSer---AspValAsnThrAsnSerSerGlyGly 257
Db       863 GGAAATGATAAACCTGTCGGAGAATCATCGGCCGGCGCTGGAGGAGAAGGTTCCGGTGGC 922

Qy       258 GlyGluSerSerSerArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGlnArg 277
Db       923 GGAGCAGAAGCAGCCGGGAAACGT---------CCGGTGGAGCTAGGCACGGCAGAGAGA 973

Qy       278 GlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeuGluGluValGluGlnArg 297
Db       974 CAAGAAATACAGATGAAGAAAGCAAAACTTAGTAACATGCTTCATGAGGTGGAGCAGAGA 1033

Qy       298 TyrArgGlnTyrHisHisGlnMetGlnIleIleValLeuSerPheGluGlnValAlaGly 317
Db      1034 TATAGACAGTACCACCAGCAGATGCAGATGGTGATCTCTTCGTTCGAGCAAGCGGCAGGG 1093

Qy       318 IleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArg 337
Db      1094 ATAGGATCAGCGAAGTCATACACGTCGCTAGCATTGAAAACCATATCAAGACAGTTCCGT 1153

Qy       338 CysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGluGlu 357
Db      1154 TGCTTGAAAGAGGCGATCGCTGGTCAGATAAAAGCGGCCAACAAGAGTCTTGGGGAGGAA 1213

Qy       358 GluGlyLeuGly------GlyLysIleGluGlySerArgLeuLysPheValAspHisHis 375
Db      1214 GATTCAGTGTCTGGTGTTGGGAGGTTTGAGGGGTCGAGGCTCAAGTTCGTGGACCACCAC 1273

Qy       376 LeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGln---------ProAsnAla 392
Db      1274 TTGAGACAGCAAAGAGCTCTTCAACAACTGGGAATGATTCAACATCCTTCCAATAATGCT 1333

Qy       393 TrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeuPhe 412
Db      1334 TGGAGACCTCAACGTGGTCTCCCAGAACGAGCCGTCTCAGTTCTCCGTGCTTGGCTCTTC 1393

Qy       413 GluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGlnThr 432
Db      1394 GAACACTTTCTTCATCCATACCCTAAGGATTCGGACAAGCACATGCTAGCTAAGCAAACA 1453

Qy       433 GlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLys 452
Db      1454 GGACTCACTCGTAGCCAGGTGTCGAACTGGTTTATAAACGCGAGAGTTCGGTTATGGAAA 1513

Qy       453 ProMetValGluGluMetTyrLeuGluGluValLysAsnGlnGluGlnAsnSerThrAsn 472
Db      1514 CCAATGGTGGAGGAGATGTACATGGAGGAAATGAAGGAGCAGGCAAAGAACATG------ 1567

Qy       473 ThrSerGlyAspAsnLysAsnLysGluThrAsnIleSerAlaProAsnGluGluLysHis 492
Db      1568 -----------GGATCCATGGAAAAGACTCCTTTGGATCAAAGCAACGAAGAT------ 1609

Qy       493 ProIleIleThrSerSerLeuLeuGlnAspGlyIleThrThrThrGlnAlaGluIleSer 512
Db      1610 ----------------------------------------TCTGCTTCAAAGTCA 1624

Qy       513 ThrSerThrIleSerThrSerProThrAlaGlyAlaSerLeuHis------HisAlaHis 530
Db      1625 ACAAGTAACCAAGAAAAGAGCCCAATGGCGGACACTAATTACCATATGAATCCCAATCAC 1684

Qy       531 AsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrThrValAspHisIle 550
Db      1685 AACGGTGACCTAGAAGGCGTCACTGGAATGCAAGGATGCCCCAAGAGACTAAGAACCAGC 1744

Qy       551 GluAsnAsnAlaLysLysGlnArgAsnAspMetHisLysPheSerProSerSerIleLeu 570
Db      1745 GACGAGACAATGATGCAGCCAATAAAT------GCGGATTTCAGCTCCAACGAGAAGCTC 1798

Qy       571 SerSerValAspMetGluAlaLysAlaArgGluSerSerAsnLysGlyPheThrAsnPro 590
Db      1799 ACGATGAAGATTCTAGAAGAACGGCAAGGGATAAGATCAGATGGTGGCTAC------CCT 1852

Qy       591 LeuMet---------AlaAlaTyrAlaMetGlyAspPheGlyArgPheAsp---ProHis 606
Db      1853 TTCATGGGTAATTTCGGGCAATACCAAATGGATGAGATGTCAAGATTTGATGTAGTCTCA 1912

Qy       607 AspGlnGlnMetThrAlaAsnPheHis------GlyAsnAsnGlyValSerLeuThrLeu 624
Db      1913 GACCAGGAGCTCATGGCGCAAAGGTACTCAGGAAACAACAATGGCGTGTCCCTCACGTTA 1972

Qy       625 GlyLeuProProSerGluAsnLeuAlaMetProValSerGln---------GlnAsnTyr 641
Db      1973 GGTTTACCTCATTGTGATAGCTTGTCGTCCACGGACCATCAGGGTTTCATGCAGACCCAC 2032

Qy       642 LeuSerAsnAspLeuGlySerArgSerGluMetGly----------------------- 653
Db      2033 CATGGGATTCCTATAGGGAGAAGAGTGAAAATAGGAGAAACAGAGGAATATGGACCCGCC 2092

Qy       654 ----------------------------------------------------SerHis 655
Db      2093 ACCATCAATGGTGGTAGCTCGACCACAACCGCACATTCATCAGCGGCAGCTGCCGCGGCT 2152

Qy       656 TyrAsnArgMetGlyTyrGluAsnIleAspPheGlnSerGlyAsnLysArgProThr 675
Db      2153 TACAATGGGATGAACATACAGAAC-----------------CAGAAGAGATATGTGGCT 2194

Qy       676 GlnLeuLeuProAspPheVal 682
Db      2195 CAGTTATTGCCCGACTTCGTT 2215
```

RESULT 11
AF406702

```
LOCUS       AF406702               2128 bp    mRNA    linear   PLN 14-JUL-2003
DEFINITION  Solanum tuberosum BEL1-related homeotic protein 29 (Bel29) mRNA,
            partial cds.
ACCESSION   AF406702
VERSION     AF406702.1  GI:22652124
KEYWORDS    .
SOURCE      Solanum tuberosum (potato)
  ORGANISM  Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum.
REFERENCE   1  (bases 1 to 2128)
  AUTHORS   Chen,H., Rosin,F.M., Prat,S. and Hannapel,D.J.
  TITLE     Interacting transcription factors from the three-amino acid loop
            extension superclass regulate tuber formation
  JOURNAL   Plant Physiol. 132 (3), 1391-1404 (2003)
   PUBMED   12857821
REFERENCE   2  (bases 1 to 2128)
  AUTHORS   Chen,H., Rosin,F.M. and Hannapel,D.J.
  TITLE     A KNOX protein of potato interacts with several members of the TALE
            family of transcription factors
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 2128)
  AUTHORS   Chen,H., Rosin,F.M. and Hannapel,D.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-AUG-2001) Department of Horticulture, Iowa State
            University, 257 Horticulture Hall, Ames, IA 50011-1100, USA
FEATURES             Location/Qualifiers
     source          1. .2128
                     /organism="Solanum tuberosum"
                     /mol_type="mRNA"
                     /cultivar="Desiree"
                     /db_xref="taxon:4113"
     gene            <1. .2128
                     /gene="Bel29"
     CDS             <1. .1704
                     /gene="Bel29"
                     /note="StBel29"
                     /codon_start=1
                     /product="BEL1-related homeotic protein 29"
```

```
                     /protein_id="AAN03626.1"
                     /db_xref="GI:22652125"
                     /translation="QGLSLSLSSSQQPGFGNFTAARELVSSPSGSASASGIQQQQQQQ
                     QSISSVPLSSKYMKAAQELLDEVVNVGKSMKSTNSTDVVVNNDVKKSKNMGDMDGQLD
                     GVGADKDGAPTTELSTGERQEIQMKKAKLVNMLDEVEQRYRHYHHQMQSVIHWLEQAA
                     GIGSAKTYTALALQTISKQFRCLKDAIIGQIRSASQTLGEEDSLGGKIEGSRLKFVDN
                     QLRQQRALQQLGMIQHNAWRPQRGLPERAVSVLRAWLFEHFLHPYPKDSDKMMLAKQT
                     GLTRSQVSNWFINARVRLWKPMVEEMYLEEIKEHEQNGLGQEKTSKLGEQNEDSTTSR
                     SIATQDKSPGSDSQNKSFVSKQDNHLPQHNPASPMPDVQRHFHTPIGMTIRNQSAGFN
                     LIGSPEIESINITQGSPKKPRNNEMLHSPNSIPSINMDVKPNEEQMSMKFGDDRQDRD
                     GFSLMGGPMNFMGGFGAYPIGEIARFSTEQFSAPYSTSGTVSLTLGLPHNENLSMSAT
                     HHSFLPIPTQNIQIGSEPNHEFGSLNTPTSAHSTSSVYETFNIQNRKRFAAPLLPDFV
                     A"
ORIGIN

Alignment Scores:
Pred. No.:                1.55e-72        Length:          2128
Score:                    1127.50         Matches:         273
Percent Similarity:       59.52%          Conservative:    77
Best Local Similarity:    46.43%          Mismatches:      150
Query Match:              31.64%          Indels:          89
DB:                       8               Gaps:            18

US-10-624-201A-2 (1-688) x AF406702 (1-2128)

Qy        161 GlyLeuSerLeuSerLeuSerProGlnLeuGlnGlnGlnIleSerPheAsnAsn------ 178
              ||||||||||||||||||||              ||||||      |||   |||
Db          4 GGGCTTTCACTTAGCCTGTCCTCG------TCCCAGCAGCCGGGGTTTGGGAACTTCACG 57

Qy        179 ------------------AsnIleSerSerSerSerProArgThrAsnAsnValThrIle 192
                               :::     |||   |||         :::  :::
Db         58 GCGGCGCGTGAGCTTGTTTCTTCGCCTTCGGGTTCGGCTTCAGCTTCAGGGATACAACAA 117

Qy        193 ArgGlyThrLeuAspGlySerSerSerAsnMetValLeuGlySerLysTyrLeuLysAla 212
              :::                |||   |||::::::   |||   |||||||||:::||||||
Db        118 CAACAACAGCAACAACAGAGTATTAGTAGTGTGCCTTTGAGTTCTAAGTACATGAAGGCT 177

Qy        213 AlaGlnGluLeuLeuAspGluValValAsnIleValGlyLysSerIleLys--------- 229
              ||||||||||||||||||||||||||||||   |||||||||||:::|||
Db        178 GCACAAGAGCTACTTGATGAAGTTGTAAAT---GTTGGAAAATCAATGAAAAGTACTAAT 234

Qy        230 ------------------GlyAspAspGlnLysLysAspAsnSerMetAsnLysGluSer 243
                               :::|||   ||||||      :::|||
Db        235 AGTACTGATGTTGTTGTTAATAATGATGTCAAGAAATCGAAGAATATG------------ 282

Qy        244 MetProLeuAlaSerAspValAsnThrAsnSerSerGlyGlyGlyGluSerSerSerArg 263
                         |||::::::          |||   |||          :::
Db        283 ------------GGCGATATGGACGGACAGTTAGACGGAGTTGGA---------GCAGAC 321

Qy        264 GlnLysAsnGluValAlaValGluLeuThrThrAlaGlnArgGlnGluLeuGlnMetLys 283
              :::                    ||||||:::|||   :::|||||||||:::|||||||||
Db        322 AAAGACGGAGCTCCAACAACTGAGCTAAGTACAGGGGAGAGACAAGAAATTCAAATGAAG 381

Qy        284 LysAlaLysLeuLeuAlaMetLeuGluGluValGluGlnArgTyrArgGlnTyrHisHis 303
              ||||||||||||:::   ||||||:::||||||||||||||||||||   |||||||||
Db        382 AAAGCAAAACTTGTTAACATGCTTGACGAGGTGGAGCAGAGGTATAGACATTATCATCAC 441

Qy        304 GlnMetGlnIleIleValLeuSerPheGluGlnValAlaGlyIleGlySerAlaLysSer 323
              |||||||||   ::::::          ||||||  ||||||||||||||||||||:::
Db        442 CAAATGCAGTCAGTGATACATTGGTTAGAGCAAGCTGCTGGCATTGGATCAGCAAAAACA 501

Qy        324 TyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArgCysLeuLysAspAlaIle 343
              ||||||   |||||||||       ||||||||||||||||||||||||||||||||||||
Db        502 TATACAGCATTGGCTTTGCAGACGATTTCGAAGCAATTTAGGTGTCTTAAGGACGCGATA 561

Qy        344 AlaGluGlnValLysAlaThrSerLysSerLeuGlyGluGluGluGlyLeuGlyGlyLys 363
                   |||::::::::::   |||::::::||||||||||||:::   ||||||||||||
Db        562 ATTGGTCAAATACGATCAGCAAGCCAGACGTTAGGCGAAGAAGATAGTTTGGGAGGGAAG 621

Qy        364 IleGluGlySerArgLeuLysPheValAspHisHisLeuArgGlnGlnArgAlaLeuGln 383
              ||||||||||||||||||||||||||||||:::   ||||||||||||||||||||||||
Db        622 ATTGAAGGTTCAAGGCTTAAATTTGTTGATAATCAGCTAAGACAGCAAAGGGCTTTGCAA 681
```

```
Qy        384 GlnIleGlyMetMetGlnProAsnAlaTrpArgProGlnArgGlyLeuProGluArgAla 403
              |||:::||||||:::|||   ||||||||||||||||||||||||||||||||||||||
Db        682 CAATTGGGAATGATCCAGCATAATGCTTGGAGACCTCAGAGAGGATTGCCCGAACGAGCT 741

Qy        404 ValSerValLeuArgAlaTrpLeuPheGluHisPheLeuHisProTyrProLysAspSer 423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        742 GTTTCTGTTCTTCGCGCTTGGCTTTTTGAACATTTCCTCCATCCTTATCCCAAGGATTCA 801

Qy        424 AspLysIleMetLeuAlaLysGlnThrGlyLeuThrArgSerGlnValSerAsnTrpPhe 443
              |||||:::||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        802 GACAAAATGATGCTAGCAAAACAAACAGGACTAACTAGGAGTCAGGTGTCGAATTGGTTC 861

Qy        444 IleAsnAlaArgValArgLeuTrpLysProMetValGluGluMetTyrLeuGluGluVal 463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db        862 ATCAATGCTCGAGTTCGTCTTTGGAAGCCAATGGTGGAAGAGATGTACTTGGAAGAGATA 921

Qy        464 LysAsnGlnGluGlnAsn--------------SerThrAsnThrSerGlyAsp---Asn 477
              |||       |||||||||                  |||:::      |||:::   |||
Db        922 AAAGAACACGAACAGAATGGGTTGGGTCAAGAAAAGACGAGCAAATTAGGTGAACAGAAC 981

Qy        478 LysAsnLysGluThrAsnIleSerAlaProAsnGluGluLysHisProIleIleThrSer 497
              ::::::      |||:::  |||          ::::::|||   |||         |||
Db        982 GAAGATTCAACAACATCAAGATCCATTGCTACACAAGACAAAAGCCCTGGTTCAGATAGC 1041

Qy        498 SerLeuLeuGlnAspGlyIleThrThrThrGlnAlaGluIleSerThrSerThrIleSer 517
                        :::     :::  |||  |||:::      ||||||
Db       1042 CAAAA-CAAGAGTTTTGTCTCAAAACAGGACAATCATTTGCCTCAACACAACCCTGCTTC 1100

Qy        518 ThrSer-ProThrAlaGlyAlaSerLeuHis--------------------------Hi 528
              |||::: |||                    |||
Db       1101 ACCAATGCCCGATGTCCAACGCCACTTCCATACCCCTATCGGTATGACCATCCGTAATCA 1160

Qy        528 sAlaHisAsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrThrValAs 548
               :::      |||:::  :::|||||      :::::::::      |||
Db       1161 GTCTGCTGGTTTCAACCTCATTGGATCACCAGAGATCGAAAGCATCAACATTACT----- 1215

Qy        548 pHisIleGluAsnAsnAlaLysLysGlnArgAsnAspMetHisLysPheSerProSerSe 568
                       :::   :::   ||||||   ||||||::      ||||||:::||
Db       1216 -------CAAGGGAGTCCAAAGAAACCGAGGAACAACGAGATGTTGCATTCACCAAACAG 1268

Qy        568 rIleLeuSerSerValAspMetGluAlaLysAlaArgGlu-------------------- 581
              ||||      |||::::::|||:::   |||       |||
Db       1269 CATT---CCATCCATCAACATGGATGTAAAGCCTAACGAGGAACAAATGTCGATGAAGTT 1325

Qy        582 ------------SerSerAsnLysGlyPheThr---------AsnProLeu-------- 591
                                ||||||:::          |||:::
Db       1326 TGGTGATGATAGGCAGGACAGAGATGGATTCTCACTAATGGGAGGACCGATGAACTTCAT 1385

Qy        592 -------MetAlaAlaTyrAlaMetGlyAspPheGlyArgPheAspProHisAspGlnGl 609
                      ||||||  :::|||:::      ||||||           :::||
Db       1386 GGGAGGATTCGGAGCCTATCCCATTGGAGAAATTGCTCGGTTTAGC------ACCGAGCA 1439

Qy        609 nMetThrAlaAsnPheHisGlyAsnAsnGlyValSerLeuThrLeuGlyLeuProProSe 629
              |   :::|||   :::      :::      ||||||||||||||||||||||   ::
Db       1440 ATTCTCAGCACCATACTCAACCAGTGGCACAGTTTCACTCACTCTTGGCCTACCACATAA 1499

Qy        629 rGluAsnLeuAlaMetProValSerGlnGlnAsnTyrLeu-------------------- 642
              :||||||||:::|||     :::      ::::::|||
Db       1500 CGAAAACCTCTCAATGTCTGCAACACACCACAGTTTCCTTCCAATTCCAACACAAAACAT 1559

Qy        643 ------------------SerAsnAspLeuGlySer---ArgSerGluMetGlySerHi 655
                               :::::::::      ||||||      :::         :::||
Db       1560 CCAAATTGGAAGTGAACCAAATCATGAGTTTGGTAGCTTAAACACACCAACATCAGCTCA 1619

Qy        655 sTyrAsnArgMetGlyTyrGluAsnIleAspPheGlnSerGlyAsnLysArgPheProTh 675
              |                ||||||      :::   |||:::      |||||||||
Db       1620 CTCAACATCAAGCGTCTATGAAACCTTCAACATTCAGAAC---AGAAAGAGGTTCGCCGC 1676

Qy        675 rGlnLeuLeuProAspPheVal 682
              |||||||||||||||||||
Db       1677 ACCCTTGTTACCAGATTTTGTT 1698

RESULT 12
```

```
AC007017
LOCUS       AC007017              120938 bp    DNA     linear   PLN 11-MAR-2002
DEFINITION  Arabidopsis thaliana chromosome 2 clone F11F19 map g6825, complete
            sequence.
ACCESSION   AC007017
VERSION     AC007017.4  GI:20197989
KEYWORDS    HTG.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 120938)
  AUTHORS   Lin,X., Kaul,S., Shea,T.P., Fujii,C.Y., Shen,M., VanAken,S.E.,
            Barnstead,M.E., Mason,T.M., Bowman,C.L., Ronning,C.M.,
            Benito,M.-I., Carrera,A.J., Creasy,T.H., Buell,C.R., Town,C.D.,
            Nierman,W.C., Fraser,C.M. and Venter,J.C.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 120938)
  AUTHORS   Lin,X.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-MAR-2000) The Institute for Genomic Research, 9712
            Medical Center Dr., Rockville, MD 20850, USA
REFERENCE   3  (bases 1 to 120938)
  AUTHORS   Town,C.D. and Kaul,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-FEB-2002) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA, cdtown@tigr.org
COMMENT     On Apr 18, 2002 this sequence version replaced gi:6598647.
            On Oct 8, 1997 this sequence version replaced gi:2443866.
             We have determined that YAC YUP8H12 is chimaeric, and is comprised
             of two distinct genomic EcoRI fragments from chromosome I. This
             submission contains the sequence from the EcoRI site at position 1

             (right end) to position 181918 of our previous Phase II
             submission YUP8H12 accession number AC000098. This fragment maps
            at
             the bottom of the right arm of the chromosome between the ATHATPAS
             and m532 markers.
FEATURES             Location/Qualifiers
     source          1. .120938
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /cultivar="Columbia"
                     /db_xref="taxon:3702"
                     /chromosome="2"
                     /map="g6825"
                     /clone="F11F19"
     misc_feature    complement(1. .7421)
                     /note="overlap with BAC clone F9C22 (AC007135:1. .7421)."
     repeat_region   complement(810. .837)
                     /rpt_family="AT_rich"
     repeat_region   complement(3770. .3798)
                     /rpt_family="AT_rich"
     repeat_region   complement(5556. .5614)
                     /rpt_family="(GA)n"
     gene            5662. .6138
                     /gene="At2g36080"
                     /note="synonym: F11F19.1"
     mRNA            <5662. .>6138
                     /gene="At2g36080"
     CDS             5662. .6138
                     /gene="At2g36080"
                     /codon_start=1
                     /product="putative DNA-binding protein RAV2"
                     /protein_id="AAM15343.1"
                     /db_xref="GI:20197990"
                     /translation="MWQQQQQQQQHQNDVVEEKEALFEKPLTPSDVGKLNRLVIPKQH
                     AERYFPLAAAAADAVEKGLLLCFEDEEGKPWRFRYSYWNSSQSYVLTKGWSRYVKEKH
                     LDAGDVVLFHRHRSDGGRFFIGWRRRGDSSSSSDSYRHVQSNASLQYYPHAGTLFI"
     repeat_region   5668. .5697
                     /rpt_family="(CAA)n"
     gene            7186. .11909
                     /gene="At2g36070"
                     /note="synonym: F11F19.2; predicted by genefinder"
     mRNA            join(<7186. .7404,9144. .9207,9294. .9396,9477. .9560,
                     9655. .9730,10165. .10539,10620. .10688,10800. .10864,
                     11096. .11168,11249. .11353,11471. .11542,11627. .11689,
                     11778. .>11909)
                     /gene="At2g36070"
     CDS             join(7186. .7404,9144. .9207,9294. .9396,9477. .9560,
                     9655. .9730,10165. .10539,10620. .10688,10800. .10864,
                     11096. .11168,11249. .11353,11471. .11542,11627. .11689,
                     11778. .11909)
                     /gene="At2g36070"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="AAM15344.1"
                     /db_xref="GI:20197991"
                     /translation="MLINCIWHIGAQAVESQRGNSKTLRLFGVNMECQLDSDWSEPST
                     PDGSNTYTTNHDQFHFYPQQQHYPPPYYMKAGSRFTDHKAASSPAIGAPTVGARLGLL
                     QGNGFASHRRFSVFSEFSKKIRGEADSNPEFQKTVKEFKERAEELQGVKEDLKVRTKQ
                     TTEKLYKQGQGVWTEAESVAKKVKESFKLGKEESAESASSSGTGTTEGEKQQQQSGST
                     EEQDTFFGKFKSSISSPKLSEAFHKPLDFAKKGLDIVKDELRGNPSKRKHLEYTPPPP
                     FTGERSTRTEMVIMPTKQSKWQKKWESLREKMQGYPVFKRLSGMSEPVVNKSQEIAED
                     VREKWETSDNPIVHKIQESFSLPDFVSEIQEAIRPVLNAYSKGDAKTLKKYCSKELIE
                     RCTAEHRAFTSQGYFFDHKLLHVSEVDIQETKMMGTTPVIIVRFQTQEIFCVRDQDGK
                     IKEGGQDTIHTVYYDWAMQQVDAAELGEDAIYPIWRLREMLRAGVQALI"
     gene            12868. .14216
                     /gene="At2g36060"
                     /note="synonym: F11F19.3; supported by full length cDNA:
                     Ceres:29951"
     mRNA            join(12868. .13060,13411. .13550,13677. .13802,13876. .14216)
                     /gene="At2g36060"
     CDS             join(13030. .13060,13411. .13550,13677. .13802,13876. .14016)
                     /gene="At2g36060"
                     /codon_start=1
                     /product="E2, ubiquitin-conjugating enzyme, putative"
                     /protein_id="AAD21451.1"
                     /db_xref="GI:4510363"
                     /translation="MTLGSGSSVVVPRNFRLLEELERGEKGIGDGTVSYGMDDGDDIY
                     MRSWTGTIIGPHNTVHEGRIYQLKLFCDKDYPEKPPTVRFHSRINMTCVNHDTGVVDS
                     KKFGVLANWQRQYTMEDILTQLKKEMAASHNRKLVQPPEGTFF"
     repeat_region   15805. .15962
                     /rpt_family="(TAAA)n"
     repeat_region   16400. .16434
                     /rpt_family="AT_rich"
     repeat_region   complement(17595. .17629)
                     /rpt_family="AT_rich"
     repeat_region   18680. .18704
                     /rpt_family="AT_rich"
     repeat_region   complement(19308. .19342)
                     /rpt_family="(CCAA)n"
     repeat_region   complement(19630. .19652)
                     /rpt_family="AT_rich"
     gene            20391. .21255
                     /gene="At2g36050"
                     /note="synonym: F11F19.32; predicted by genefinder;
                     supported by cDNA: gi_14994264_gb_AY044326 1_"
     mRNA            <20391. .21255
                     /gene="At2g36050"
     CDS             20391. .21176
                     /gene="At2g36050"
                     /codon_start=1
                     /product="expressed protein"
                     /protein_id="AAD21452.1"
                     /db_xref="GI:4510364"
                     /translation="MKLPFLNKNHSTSSYSSNSSSSSWPWPSCNQNPKTLSFRATITF
                     TNPIHDQDDDELDLLDPPEITDSVENVIKGLRSSERLIFESKGETNSILEEATSKREE
                     EDEEEGFMLFSLESDDPYSDFKRSMEEMVEAHALHHDWKSLEKLLLQFLKVNAKTSHR
                     YIFAAFVDLLMNLALDTKKAIINNDISKDDGVSASRAAAAGEASTSCCNSMTLGESPS
                     SPLSFYTSCSSSSSSDETSSMSVRFLPLSSLLEMDEKTKEILV"
     repeat_region   21508. .21543
                     /rpt_family="AT_rich"
     repeat_region   22287. .22365
                     /rpt_family="AT_rich"
```

```
     repeat_region   22577. .22612
                     /rpt_family="AT_rich"
     repeat_region   23077. .23102
                     /rpt_family="AT_rich"
     repeat_region   complement(25817. .25860)
                     /rpt_family="AT_rich"
     repeat_region   complement(25894. .25930)
                     /rpt_family="AT_rich"
     gene            26200. .26382
                     /gene="At2g36040"
                     /note="synonym: F11F19.5; predicted by genefinder"
     mRNA            <26200. .>26382
                     /gene="At2g36040"
     CDS             26200. .26382
                     /gene="At2g36040"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="AAD21453.1"
                     /db_xref="GI:4510365"
                     /translation="MDASDSNMKDGFNVEEEFQKVEEQVSKTYEEKKTFKVRSIDGDK
                     LGFALSESSRSESDEE"
     repeat_region   26564. .26660
                     /rpt_family="AT_rich"
     repeat_region   complement(28601. .28628)
                     /rpt_family="(TA)n"
     gene            29257. .29469
                     /gene="At2g36030"
                     /note="synonym: F11F19.6; predicted by genefinder"
     mRNA            <29257. .>29469
                     /gene="At2g36030"
     CDS             29257. .29469
                     /gene="At2g36030"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="AAD21454.1"
                     /db_xref="GI:4510366"
                     /translation="MTVDEWRLLDVSTAFGSGENSVPDFGESADDDAVVVWLVKGEEA
                     VEKLSEIVFGVNLVLENHLKHGLSEI"
     repeat_region   29525. .29605
                     /rpt_family="(GGA)n"
     gene            31440. .33782
                     /gene="At2g36020"
                     /note="synonym: F11F19.7; predicted by
                     genefinder;supported by full length cDNA: Ceres:13310"
     mRNA            join(31440. .31565,31654. .31773,31864. .31919,32009. .32107,
                     32534. .32590,32698. .32761,32865. .32999,33109. .33192,
                     33299. .33782)
                     /gene="At2g36020"
     CDS             join(31888. .31919,32009. .32107,32534. .32590,32698. .32761,
                     32865. .32999,33109. .33192,33299. .33604)
                     /gene="At2g36020"
                     /codon_start=1
                     /product="expressed protein"
                     /protein_id="AAD21455.2"
                     /db_xref="GI:20197992"

Alignment Scores:
Pred. No.:              6.63e-61        Length:         120938
Score:                  999.00          Matches:        305
Percent Similarity:     42.53%          Conservative:   82
Best Local Similarity:  33.52%          Mismatches:     201
Query Match:            28.04%          Indels:         323
DB:                     8               Gaps:           31

US-10-624-201A-2 (1-688) x AC007017 (1-120938)

Qy         22 HisGlyAsn-----------------SerAsnAsnAsnAsnIleGlnThrLeuTyrLeu 35
              |||||||||                 ::::::      :::|||||||||   |||
Db      65344 CACGGAAACCCACCGGAGATCTCTGCCGGATCCGACGGTGGTCTTCAAACGTTGATCCTC 65403

Qy         36 MetAsnProAsnAsnTyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGlnGlnGln 55
              |||||||||       |||:::|||   ||||||      |||
Db      65404 ATGAATCCAACTACTTACGTTCAG---TACACCCAACAAGACAACGACTCGAACAACAAC 65460

Qy         56 LeuLeuPheLeuAsnSerSerProAlaAlaSerAsnAlaLeuCysHisAlaAsnIle--- 74
                         |||:::|||      ::::::|||         :::   |||
Db      65461 ------------AACAACAGCAACAATAGCAACAACAACAACACAAACACAAACACAAAC 65508

Qy         75 ----------------------------GlnHisAlaProLeu------GlnGlnGln 82
                                           |||||||||                |||
Db      65509 AACAACAACAGTAGTTTCGTTTTCCTCGATTCCCACGCGCCGCAGCCAAACGCGAGCCAG 65568

Qy         83 HisPheValGlyValProLeu--------------ProAlaValSerLeuHisAspGln 97
              |||||||||:::||||||                :::::::::      |||
Db      65569 CAGTTCGTCGGAATACCACTCTCAGGTCACGAAGCTGCTTCCATTACAGCCGCCGACAAC 65628

Qy         98 IleAsn---HisHisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnSerGlnGln 116
              |||:::      ||||||         |||:::      :::         |||   |||
Db      65629 ATCTCCGTACTTCACGGTTATCCTCCGCGCGTGCAGTACAGTCTCTACGGTAGCCACCAA 65688

Qy        117 ValIleValProSerSerThrGlyValSerAlaThrSerCysGlyGlyIleThrThrAsp 136
              |||
Db      65689 GTG------------------------------------------------------- 65691

Qy        137 LeuAlaSerGlnLeuAlaPheGlnArgProIleProThrProGlnHisArgGln------ 154
                                                 |||   |||:::|||
Db      65692 ----------------------------------GATCCCACTCACCAGCAAGCCGCG 65715

Qy        155 ------GlnGlnGlnGlnGlnGlyGlyLeuSerLeuSerLeuSerProGlnLeuGlnGln 172
                    :::   |||   |||||||||||:::||||||   |||   ||||||
Db      65716 TGTGAGACGCCACGCGCGCAGCAAGGCCTCTCTTTAACCCTCTCGTCTCAACAGCAGCAG 65775

Qy        173 GlnIleSerPheAsnAsn--------------AsnIleSerSerSerSerProArgThr 187
              |||          :::                ::::::         |||
Db      65776 CAACAGCAACATCATCAACAACACCAGCCTATTCACGTCGGATTCGGGTCCGGACATGGA 65835

Qy        188 AsnAsnValThrIleArgGlyThrLeuAspGlySerSer--------------SerAsn 202
                 ::::::   :::                |||||              :::|||
Db      65836 GAAGATATCCGGGTCGGGTCTGGCTCTACAGGATCGGGGGTAACAAACGGTATAGCTAAT 65895

Qy        203 MetValLeuGlySerLysTyrLeuLysAlaAlaGlnGluLeuLeuAspGluValValAsn 222
              :::|||      ||||||||||||||||||||||||||||||||||||||||||||||
Db      65896 CTTGTT---AGCTCCAAGTACTTGAAGGCAGCACAAGAGCTTCTTGACGAAGTAGTCAAC 65952

Qy        223 IleValGlyLysSerIleLysGlyAspAspGln-----------LysLysAspAsnSer 238
                              :::             |||          ||||||   :::
Db      65953 GCTGATTCCGATGACATGAACGCTAAATCCCAACTATTCTCATCGAAAAAGGGTAGTTGC 66012

Qy        239 MetAsnLysGluSerMetProLeuAlaSer---AspValAsnThrAsnSerSerGlyGly 257
              |||   :::   :::       :::|||                     |||||||||
Db      66013 GGAAATGATAAACCTGTCGGAGAATCATCGGCCGGCGCTGGAGGAGAAGGTTCCGGTGGC 66072

Qy        258 GlyGluSerSerSerArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGlnArg 277
              |||       ::::::   ::::::            |||||||||   ||||||::|||
Db      66073 GGAGCAGAAGCAGCCGGGAAACGT---------CCGGTGGAGCTAGGCACGGCAGAGAGA 66123

Qy        278 GlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeu----------------- 291
              ||||||:::||||||||||||||||||||||       |||
Db      66124 CAAGAAATACAGATGAAGAAAGCAAAACTTAGTAACATGCT-TCATGAGGTAACAACAAT 66182

Qy        291 ------------------------------------------------------------ 291

Db      66183 ACGTATTATATACAATAAGCCTACATATACTGAGTGCTCTCTTATATCACACATGAAGTG 66242

Qy        291 ------------------------------------------------------------ 291

Db      66243 GACTTAAAAGTTCTGTCCCTAAGATTTTCAGTCATTTTCACAATTCATTTTATATGTATA 66302

Qy        291 ------------------------------------------------------------ 291

Db      66303 TTAAGCTATCCTTTTTTTTTTTGCTCTTCCGATTTCTCAAATTGACAACTTACAAGATT 66362

Qy        291 ------------------------------------------------------------ 291

Db      66363 GATGAAATCTTTTTTTTTTTGTCCCTCTCTTAGTAAAGACTAACAACTTTTATAGATTAT 66422
```

```
Qy        291 ------------------------------------------------------------ 291

Db      66423 ATTTATGTAGACCGTACACTTATGTTCATAAGACTCACGTTACTTTGTTGGTTGAAAAAT 66482

Qy        292 GluGluValGluGlnArgTyrArgGlnTyrHisHisGlnMetGlnIleIleValLeuSer 311
              |||:::|||||||||||||||||||||||||||||   |||||||||:::::::::   |||
Db      66483 GAAAAGGTGGAGCAGAGATATAGACAGTACCACCAGCAGATGCAGATGGTGATCTCTTCG 66542

Qy        312 PheGluGlnValAlaGlyIleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAla 331
              |||||||||   |||||||||||||||||||||||||||||   |||||||||
Db      66543 TTCGAGCAAGCGGCAGGGATAGGATCAGCGAAGTCATACACGTCGCTAGCATTGAAAACC 66602

Qy        332 IleSerLysGlnPheArgCysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSer 351
              ||||||:::|||||||||||||||||||:::|||||||||   |||:::||||||   :::
Db      66603 ATATCAAGACAGTTCCGTTGCTTGAAAGAGGCGATCGCTGGTCAGATAAAAGCGGCCAAC 66662

Qy        352 LysSerLeuGlyGluGluGluGlyLeuGly------GlyLysIleGluGlySerArgLeu 369
              |||||||||||||||||:::   :::         |||:::   |||||||||||||||
Db      66663 AAGAGTCTTGGGGAGGAAGATTCAGTGTCTGGTGTTGGGAGGTTTGAGGGGTCGAGGCTC 66722

Qy        370 LysPheValAspHisHisLeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGln 389
              |||||||||||||||||||||||||||||||||||||||||||||:::||||||:::|||
Db      66723 AAGTTCGTGGACCACCACTTGAGACAGCAAAGAGCTCTTCAACAACTGGGAATGATTCAA 66782

Qy        390 ---------ProAsnAlaTrpArgProGlnArgGlyLeuProGluArgAlaValSerVal 406
                      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      66783 CATCCTTCCAATAATGCTTGGAGACCTCAACGTGGTCTCCCAGAACGAGCCGTCTCAGTT 66842

Qy        407 LeuArgAlaTrpLeuPheGluHisPheLeuHisPro---------------------- 418
              |||||||||||||||||||||||||||||||||
Db      66843 CTCCGTGCTTGGCTCTTCGAACACTTTCTTCATCCGTAAGTAACCCACTATACACACATG 66902

Qy        418 ------------------------------------------------------------ 418

Db      66903 CACTACTTTTCATCACTGGAACACAAACTCAGTCTCTAAGTTGTCTTTCTTTTAATTAGA 66962

Qy        419 TyrProLysAspSerAspLysIleMetLeuAlaLysGlnThrGlyLeuThrArgSerGln 438
              |||||||||||||||||||   ||||||||||||||||||||||||||||||||||||||
Db      66963 TACCCTAAGGATTCGGACAAGCACATGCTAGCTAAGCAAACAGGACTCACTCGTAGCCAG 67022

Qy        438 ------------------------------------------------------------ 438

Db      67023 GTACGTTTCAAAATGATTCTTTCTTGTGTTATTAGGTTACTTAAGAAAGCGTTTTTAAGT 67082

Qy        439 ------------------------ValSerAsnTrpPheIleAsnAlaArgValArgLe 450
                                       |||||||||||||||||||||||||||||||||
Db      67083 AAATTTATTTTCTTTAATTATTTAGGTGTCGAACTGGTTTATAAACGCGAGAGTTCGGTT 67142

Qy        450 uTrpLysProMetValGluGluMetTyrLeuGluGluValLysAsnGlnGluGlnAsnSe 470
              |||||||||||||||||||||||||||:::||||||:::|||   |||   :::|||
Db      67143 ATGGAAACCAATGGTGGAGGAGATGTACATGGAGGAAATGAAGGAGCAGGCAAAGAACAT 67202

Qy        470 rThrAsnThrSerGlyAspAsnLysAsnLysGluThrAsnIleSerAlaProAsnGluGl 490
                                     ::::::|||   :::         ||||||::
Db      67203 G-----------------GGATCCATGGAAAAGACTCCTTTGGATCAAAGCAACGAAGA 67244

Qy        490 uLysHisProIleIleThrSerSerLeuLeuGlnAspGlyIleThrThrThrGlnAlaGl 510
              :                                                      |||
Db      67245 T-------------------------------------------------TCTGCTTC 67253

Qy        510 uIleSerThrSerThrIleSerThrSerProThrAlaGlyAlaSerLeuHis------Hi 528
              |||||||||           ||||||   |||      :::   |||
Db      67254 AAAGTCAACAAGTAACCAAGAAAAGAGCCCAATGGCGGACACTAATTACCATATGAATCC 67313

Qy        528 sAlaHisAsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrThrValAs 548
              ||||||           |||         |||      :::        :::
Db      67314 CAATCACAACGGTGACCTAGAAGGCGTCACTGGAATGCAAGGAAGCCCCAAGAGACTAAG 67373

Qy        548 pHisIleGluAsnAsnAlaLysLysGlnArgAsnAspMetHisLysPheSerProSerSe 568
                  :::         :::         |||                |||||   :::
Db      67374 AACCAGCGACGAGACAATGATGCAGCCAATAAAT------GCGGATTTCAGCTCCAACGA 67427

Qy        568 rIleLeuSerSerValAspMetGluAlaLysAlaArgGluSerSerAsnLysGlyPheTh 588
```

```
                  |||:::         :::|||   :::            |||:::   |||:::
Db      67428 GAAGCTCACGATGAAGATTCTAGAAGAACGGCAAGGGATAAGATCAGATGGTGGCTAC-- 67485

Qy        588 rAsnProLeuMet---------AlaAlaTyrAlaMetGlyAspPheGlyArgPheAsp-- 604
                  |||   |||                |||   |||   :::      |||||||
Db      67486 ----CCTTTCATGGGTAATTTCGGGCAATACCAAATGGATGAGATGTCAAGATTTGATGT 67541

Qy        605 -ProHisAspGlnGlnMetThrAlaAsnPheHis------GlyAsnAsnGlyValSerLe 622
                   |||||::::::   |||       :::          ||||||||||||||||
Db      67542 AGTCTCAGACCAGGAGCTCATGGCGCAAAGGTACTCAGGAAACAACAATGGCGTGTCCCT 67601

Qy        622 uThrLeuGlyLeuProProSerGluAsnLeuAlaMetProValSerGln---------Gl 639
              ||||||||||||||||       ::::::|||:::            |||         ||
Db      67602 CACGTTAGGTTTACCTCATTGTGATAGCTTGTCGTCCACGCACCATCAGGGTTTCATGCA 67661

Qy        639 nAsnTyrLeuSerAsnAspLeuGlySerArgSerGluMetGly----------------- 653
              |   :::            :::|||   |||   ::::::|||
Db      67662 GACCCACCATGGGATTCCTATAGGGAGAAGAGTGAAAATAGGAGAAACAGAGGAATATGG 67721

Qy        653 ------------------------------------------------------------ 653

Db      67722 ACCCGCCACCATCAATGGTGGTAGCTCGACCACAACCGCACATTCATCAGCGGCAGCTGC 67781

Qy        654 -SerHisTyrAsnArgMetGlyTyrGluAsnIleAspPheGlnSerGlyAsnLysArgPh 673
               :::   ||||||   |||      :::|||                      ||||||::
Db      67782 CGCGGCTTACAATGGGATGAACATACAGAAC-----------------CAGAAGAGATA 67823

Qy        673 eProThrGlnLeuLeuProAspPheVal 682
              :      |||||||||||||||||||
Db      67824 TGTGGCTCAGTTATTGCCCGACTTCGTT 67851
```

```
RESULT 13
AK120844
LOCUS       AK120844               2489 bp    mRNA    linear   PLN 29-OCT-2003
DEFINITION  Oryza sativa (japonica cultivar-group) cDNA clone:J023022M14, full
            insert sequence.
ACCESSION   AK120844
VERSION     AK120844.1  GI:37990467
KEYWORDS    FLI_CDNA; CAP trapper.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1
  AUTHORS   The Rice Full-Length cDNA Consortium, National Institute of
            Agrobiological Sciences Rice Full-Length cDNA Project Team:,
            Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,
            Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,
            Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,
            Ohtsuki,K., Shishiki,T., Foundation of Advancement of International
            Science Genome Sequencing & Analysis Group:, Otomo,Y., Murakami,K.,
            Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,
            Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,
            Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J.,
            Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,
            Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN:,
            Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,
            Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,
            Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y.,
            Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
            Yoshino,M. and Hayashizaki,Y.
  TITLE     Collection, mapping, and annotation of over 28,000 cDNA clones from
            japonica rice
  JOURNAL   Science 301 (5631), 376-379 (2003)
  MEDLINE   22752273
   PUBMED   12869764
REFERENCE   2
  AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,
            Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W.,
            Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T.,
            Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K.,
            Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I.,
```

Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M., Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M., Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M., Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A., Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakahama,Y., Nakamura,M., Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K., Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H., Osato,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K., Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S., Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W., Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and Yoshimura,A.

TITLE Collection, mapping, and annotation of 28K full-length cDNA clones from japonica rice

JOURNAL Unpublished

REFERENCE 3 (bases 1 to 2489)

AUTHORS Kikuchi,S.

TITLE Direct Submission

JOURNAL Submitted (31-JAN-2003) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)

COMMENT This clone is one of the 32K full-length cDNA clones from japonica rice.
URL : http://cdna01.dna.affrc.go.jp/cDNA/
NIAS Rice Full-Length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., Yamamoto,M. and Nakahama,Y.
FAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y., Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M., Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J., Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S., Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S., Yoshimura,A., Matsubara,K. and Murakami,K.
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J., Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N., Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K., Yasunishi,A. and Hayashizaki,Y.

FEATURES Location/Qualifiers

source 1. .2489
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J023022M14"

ORIGIN

Alignment Scores:

Pred. No.:	8.41e-61	Length:	2489
Score:	967.50	Matches:	283
Percent Similarity:	47.42%	Conservative:	76
Best Local Similarity:	37.38%	Mismatches:	215
Query Match:	27.15%	Indels:	183
DB:	8	Gaps:	33

US-10-624-201A-2 (1-688) x AK120844 (1-2489)

```
Qy          2 TyrTyrGlnGlyThrSerAspAsnThrAsn---IleGlnAlaAspHisGlnGlnArgHis 20
              |||:::   |||           |||:::   :::||||||
Db        343 TACTTCACAGGCGGCGGCGCCGGCACCGACGTCGTCCAGGCC----------------- 384

Qy         21 AsnHisGlyAsnSerAsnAsnAsnAsnIleGlnThrLeuTyrLeuMetAsnProAsnAsn 40
                     :::     :::   :::||||||||||||||||||||||||   :::
Db        385 -----------GCCGGCACCGACGGCCTGCAGACGCTCTACCTCATGAACCCC---AGC 429

Qy         41 TyrMetGlnGlyTyrThrThrSerAspThr--------------GlnGlnGlnGlnGln 55
              |||:::   |||:::|||   :::
Db        430 TACGTC---GGCTTCACGGACGCCGCCGCGGCGCCCGGCGGCGGCGCCGCCGCCGCCAAC 486

Qy         56 LeuLeuPheLeuAsnSerSerProAlaAlaSerAsnAlaLeuCysHisAlaAsnIleGln 75
              ::::::|||||||||||          |||   :::   |||      |||:::
Db        487 ATGGTGTTCCTCAACTCG---------GCGGTGAGCACCCTCACGCCGGCGAGCTTCAGC 537

Qy         76 His------------AlaProLeuGlnGlnGlnHisPheValGlyValProLeuProAla 91
              |||            |||         |||||||||||||:::||||||   :::
Db        538 CACCACCACCAGCCGACACCG---GCGGCGCAGCACTTCGTCGGCATCCCTCTCCAGTCG 594

Qy         92 ValSerLeuHisAspGlnIleAsnHisHisGlyLeuLeuGlnArgMetTrpAsnAsnGln 111
                                                          :::|||
Db        595 GGG-----------------------------------TACAACCTGTGGGGC------ 612

Qy        112 AspGlnSerGlnGlnValIleValProSerSerThrGlyValSerAlaThrSerCysGly 131
                                      |||   :::||||||   :::       |||
Db        613 ----------------------CCTGACGCCACCGGCGGCAACGACGTGTCG------ 642

Qy        132 GlyIleThrThrAspLeuAlaSerGlnLeuAlaPheGlnArgProIleProThrProGln 151
                                                                   |||:::
Db        643 ------------------------------------------------------CCGCCACGC 651

Qy        152 HisArgGlnGlnGlnGln---------GlnGlnGlyGlyLeuSerLeuSerLeuSerPro 168
              |||      |||||||||           |||      :::   ::::::||||||
Db        652 CACGGCGCCCAGCAGCAAGCGCCAGCAGCAGCTGGCACGTCGGCGGCGGCGGTCAGCCCC 711

Qy        169 GlnLeuGlnGlnGlnIleSerPheAsnAsnAsnIleSerSerSerSerProArgThrAsn 188
                |||                                    ||||||
Db        712 GTGCTGAGC-------------------------CTGTCGTCGCGTGAGGCGGCGCCG 744

Qy        189 AsnValThrIleArgGlyThrLeuAspGlySerSer--------------------Ser 201
              ||||||:::                 ::::::
Db        745 CCGGTCACCGTGGCCGCCGCCGCCGCCGCCGCCGTCCCCGGCGGCACGGACCAGGAG 804

Qy        202 AsnMetValLeuGlySerLysTyrLeuLysAlaAlaGlnGluLeuLeuAspGluValVal 221
                 :::|||:::   |||:::|||||||||||||||||||||||||||||||   |||
Db        805 AAGGTGGTGATGCGCTCCAGGTACCTGAAGGCGGCGCAGGAGCTGCTCGACGAGGCGGTG 864

Qy        222 AsnIleValGlyLysSerIleLysGlyAspAspGlnLysLysAspAsnSerMetAsnLys 241
                 :::   |||   |||                 :::||||||::::::|||
Db        865 AGC---GTCAGCAAGGGCGCCGCGACGGCGGTGAAGAAGAAGGAGGACTCGGAGGGCGGC 921

Qy        242 GluSerMetProLeuAlaSerAspValAsnThrAsnSerSerGlyGlyGlyGluSerSer 261
              |||                                      ||||||   |||
Db        922 GTGTCC-------------------------GGTGGCGGCGGCGGTGCGGAGGACGGT 954

Qy        262 SerArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGlnArgGlnGluLeuGln 281
                    |||:::      |||   |||:::::||||||:::|||||||||||||||||
Db        955 GGCGGCAGCAAGAGCGGCGCCGCCGCCGAGATGTCGACGGCGGAGCGGCAGGAGCTGCAG 1014

Qy        282 MetLysLysAlaLysLeuLeuAlaMetLeuGluGluValGluGlnArgTyrArgGlnTyr 301
              |||||||||:::|||||||||   |||||||:::||||||||||||||||||||||||||
Db       1015 ATGAAGAAGAGCAAGCTACTGAACATGCTCGATGAGGTGGAGCAGCGGTACAGGCAGTAC 1074

Qy        302 HisHisGlnMetGlnIleIleValLeuSerPheGluGlnValAlaGlyIleGlySerAla 321
              |||   |||||||||   :::      :::||||||      ||||||   |||||||||
Db       1075 CACCGGCAGATGCAGGGGGTGGCGGCGGCGTTCGAGGCGGCGGCGGGGGCCGGGTCGGCG 1134

Qy        322 LysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArgCysLeuLysAsp 341
                 :::||||||   |||||||||       ||||||:::|||||||||||||||:::|||
Db       1135 ACGACGTACACGTCGCTGGCGCTGCGCACCATCTCGCGGCAGTTCCGGTGCCTCCGCGAC 1194
```

```
Qy        342 AlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGlu-----------Glu 357
              |||||||||   ||||||:::|||   |||:::   |||||||||           :::
Db       1195 GCCATCGCGGCGCAGGTCAGGGCGGCGAGCCGGGGGCTCGGGGAGGACTGCGGCGACGAC 1254

Qy        358 GluGlyLeuGlyGly---------LysIleGluGlySerArgLeuLysPheValAspHis 374
              ||||||   ||||||                    |||||||||||:::|||:::|||||||
Db       1255 GAGGGCGGCGGTGGCGGCGGGAGGACGACGGTGGGGTCGCGGCTGCGGTTCATCGACCAC 1314

Qy        375 HisLeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGlnProAsnAla------ 392
              ||||||||||||||||||:::||||||:::||||||:::          |||
Db       1315 CAGCTCCGGCAGCAGCGCGCGATGCAGCAGCTTGGCATGGTGCACGCCGCCGCCGCCGGT 1374

Qy        393 ------------------TrpArgProGlnArgGlyLeuProGluArgAlaValSerVal 406
                               ||||||||||||||||||||||||||||||||||||||||||
Db       1375 GGCGCCGCCGGCGGCGGGTGGCGGCCGCAGCGTGGCCTCCCCGAGCGCGCCGTCTCCGTC 1434

Qy        407 LeuArgAlaTrpLeuPheGluHisPheLeuHisProTyrProLysAspSerAspLysIle 426
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db       1435 CTCCGCGCTTGGCTCTTCGAGCATTTCCTCCACCCATACCCCAAGGATTCGGACAAGGTC 1494

Qy        427 MetLeuAlaLysGlnThrGlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAla 446
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1495 ATGCTCGCCAAGCAAACCGGCCTCACCAGGAGCCAGGTGTCGAACTGGTTCATCAACGCG 1554

Qy        447 ArgValArgLeuTrpLysProMetValGluGluMetTyrLeuGluGluValLysAsnGln 466
              |||||||||||||||||||||||||||||||||||||    ||||||    |||   :::
Db       1555 AGGGTGAGGCTGTGGAAGCCAATGGTGGAGGAGATGTACGCCGAGGAGACCAAGGCCAAG 1614

Qy        467 GluGlnAsnSerThrAsnThrSerGlyAspAsnLysAsnLysGluThrAsnIleSerAla 486
              |||:::
Db       1615 GAGGAG------------------------------------------------------ 1620

Qy        487 ProAsnGluGluLysHisProIleIleThrSerSerLeuLeuGlnAspGlyIleThrThr 506
                    ||||||:::|||           ::::::           |||
Db       1621 ---GAGGAGGAGGAGCACGACGCGGCGGCGGCCGCCGCCGGCGACCGTGGCGGCGTGGCG 1677

Qy        507 ThrGlnAla-----------------GluIleSerThrSerThrIleSerThrSerPro 520
              ||||||                                          |||:::||||||
Db       1678 GAGCAGGCTCCGAGTAAGCCGGACGACAGTGCTGGCATTGGCATGTCGTCGTCGTCGCCG 1737

Qy        521 ThrAlaGlyAlaSer------------LeuHis------------HisAlaHisAsnPhe 532
                    ||||||           :::|||           ||||||
Db       1738 GCGGTGGCGGCGTCCAGGTCAGTTGGTGTGCACGCCGGAGACCAGCACGCTCAGGCGAGC 1797

Qy        533 SerPheLeuGlySerPheAsnMetAspAsn--------------------------Thr 543
                 :::   |||             |||:::
Db       1798 TTCTACGGCGGCGGCGGCGGCGGCGACGATCCGTTCCAGTGCAGGATCAAGAAGGCGAGG 1857

Qy        544 ThrThrThrValAsp-----------------------HisIleGluAsnAsnAlaLys 555
              |||||||||   |||                       |||:::         |||
Db       1858 ACGACGACGGCGGACGAGCCGGCGGCGGCCGCGGCGTTCCACGTCTCCGGCGAGGCCGCC 1917

Qy        556 ---LysGlnArgAsnAspMetHisLysPheSerProSerSer------------------ 568
                  |||      :::   ||||||:::   :::
Db       1918 GTGAGCCACCGGGAGCTGCTAATGAAGTTCACTGAGGCCGGCGGCGAGGGCGTGAGAACG 1977

Qy        569 ---------IleLeuSerSerValAspMetGluAlaLysAlaArgGluSerSerAsnLys 585
                      ::::::       |||   :::
Db       1978 GGCCACCCGCACGTCAACGACGACGACGACGACGTCCCTGGCGGCGCCGGCGGGTACTCG 2037

Qy        586 GlyPheThrAsnProLeuMetAlaAlaTyrAlaMetGlyAspPheGlyArgPheAspPro 605
              ||||||              ||||||                ||||||
Db       2038 CTGTTCACG-----------GCGGCGCAGTACGGGCATCAGTTCGGC------TCGGAC 2079

Qy        606 HisAspGlnGlnMetThrAlaAsnPheHisGlyAsnAsnGlyValSerLeuThrLeuGly 625
              |||                  |||      ||||||||||||||||||||
Db       2080 CATTTCGCGTTCGCCGGGCACGGCGGCGGAGGCGGCGGCGGCGTGTCGCTCACGCTCGGC 2139

Qy        626 LeuProProSerGluAsnLeuAlaMetProValSerGlnGlnAsnTyrLeuSerAsnAsp 645
              ||||||         :::         |||   |||         :::|||
Db       2140 CTCCCGCACGGCGCCGAC---CAGACGCCGGCGTCG---------TTTCTCATAGGCGCC 2187
```

```
Qy        646 LeuGlySerArgSerGluMetGlySerHisTyrAsnArgMetGlyTyrGluAsnIleAsp 665
                                       |||:::               ||||||:::   ::::::
Db       2188 GGCGCCGGCAGCGACGGCGGCGGCGCGCCGGTGACGACGGCCGGCTACGAC---ATGAAC 2244

Qy        666 PheGlnSerGlyAsnLysArgPheProThrGlnLeuLeuProAspPheVal 682
                    ||||||      |||           ||||||:::   |||||||||
Db       2245 ATGCAGAGC---ACCAAGTCGTTGGCTGCTCAGCTCATGAGAGACTTCGTG 2292

RESULT 14
AK120813
LOCUS       AK120813               2643 bp    mRNA    linear   PLN 29-OCT-2003
DEFINITION  Oryza sativa (japonica cultivar-group) cDNA clone:J023014M03, full
            insert sequence.
ACCESSION   AK120813
VERSION     AK120813.1  GI:37990436
KEYWORDS    FLI_CDNA; CAP trapper.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1
  AUTHORS   The Rice Full-Length cDNA Consortium, National Institute of
            Agrobiological Sciences Rice Full-Length cDNA Project Team:,
            Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,
            Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,
            Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,
            Ohtsuki,K., Shishiki,T., Foundation of Advancement of International
            Science Genome Sequencing & Analysis Group:, Otomo,Y., Murakami,K.,
            Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,
            Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,
            Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J.,
            Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,
            Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN:,
            Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,
            Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,
            Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y.,
            Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
            Yoshino,M. and Hayashizaki,Y.
  TITLE     Collection, mapping, and annotation of over 28,000 cDNA clones from
            japonica rice
  JOURNAL   Science 301 (5631), 376-379 (2003)
  MEDLINE   22752273
   PUBMED   12869764
REFERENCE   2
  AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,
            Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W.,
            Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T.,
            Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K.,
            Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I.,
            Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M.,
            Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M.,
            Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M.,
            Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A.,
            Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakahama,Y.,
            Nakamura,M., Namiki,T., Narikawa,R., Niikura,J., Nishi,K.,
            Nomura,K., Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M.,
            Ooka,H., Osato,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C.,
            Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K.,
            Shibata,K., Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y.,
            Sugano,S., Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M.,
            Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
            Tanaka,T., Tomaru,A., Toya,T., Tsunoda,Y., Ueda,M., Waki,K.,
            Xie,Q., Yahagi,W., Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J.,
            Yokomizo,S. and Yoshimura,A.
  TITLE     Collection, mapping, and annotation of 28K full-length cDNA clones
            from japonica rice
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 2643)
  AUTHORS   Kikuchi,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-JAN-2003) Shoshi Kikuchi, National Institute of
```

Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)

COMMENT This clone is one of the 32K full-length cDNA clones from japonica rice.
URL : http://cdna01.dna.affrc.go.jp/cDNA/
NIAS Rice Full-Length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., Yamamoto,M. and Nakahama,Y.
FAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y., Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M., Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J., Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S., Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S., Yoshimura,A., Matsubara,K. and Murakami,K.
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J., Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N., Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K., Yasunishi,A. and Hayashizaki,Y.

FEATURES Location/Qualifiers
source 1. .2643
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J023014M03"

ORIGIN

Alignment Scores:

Pred. No.:	3.34e-59	Length:	2643
Score:	946.00	Matches:	283
Percent Similarity:	46.03%	Conservative:	76
Best Local Similarity:	36.28%	Mismatches:	215
Query Match:	26.55%	Indels:	206
DB:	8	Gaps:	34

US-10-624-201A-2 (1-688) x AK120813 (1-2643)

```
Qy          2 TyrTyrGlnGlyThrSerAspAsnThrAsn---IleGlnAlaAspHisGlnGlnArgHis 20
              |||:::   |||             |||:::   :::||||||
Db        343 TACTTCACAGGCGGCGGCGCCGGCACCGACGTCGTCCAGGCC----------------- 384

Qy         21 AsnHisGlyAsnSerAsnAsnAsnAsnIleGlnThrLeuTyrLeuMetAsnProAsnAsn 40
                        :::      :::   :::||||||||||||||||||||||||   :::
Db        385 ------------GCCGGCACCGACGGCCTGCAGACGCTCTACCTCATGAACCCC---AGC 429

Qy         41 TyrMetGlnGlyTyrThrThrSerAspThr--------------GlnGlnGlnGlnGln 55
              |||:::   |||:::|||   :::
Db        430 TACGTC---GGCTTCACGGACGCCGCCGCGGCGCCCGGCCGGCGGCGCCGCCGCCGCCAAC 486

Qy         56 LeuLeuPheLeuAsnSerSerProAlaAlaSerAsnAlaLeuCysHisAlaAsnIleGln 75
              ::::::||||||||||||          |||   :::   |||       |||:::
Db        487 ATGGTGTTCCTCAACTCG---------GCGGTGAGCACCCTCACGCCGGCGAGCTTCAGC 537

Qy         76 His------------AlaProLeuGlnGlnGlnHisPheValGlyValProLeuProAla 91
              |||             |||         ||||||||||||||:::||||||   :::
Db        538 CACCACCACCAGCCGACACCG---GCGGCGCAGCACTTCGTCGGCATCCCTCTCCAGTCG 594

Qy         92 ValSerLeuHisAspGlnIleAsnHisHisGlyLeuLeuGlnArgMetTrpAsnAsnGln 111
                                                                  :::|||
Db        595 GGG---------------------------------TACAACCTGTGGGGC------ 612

Qy        112 AspGlnSerGlnGlnValIleValProSerSerThrGlyValSerAlaThrSerCysGly 131
                                      |||   :::||||||   :::      |||
Db        613 ----------------------CCTGACGCCACCGGCGGCAACGACGTGTCG------ 642

Qy        132 GlyIleThrThrAspLeuAlaSerGlnLeuAlaPheGlnArgProIleProThrProGln 151
                                                                  |||:::
Db        643 ----------------------------------------------------CCGCCACGC 651

Qy        152 HisArgGlnGlnGlnGln---------GlnGlnGlyGlyLeuSerLeuSerLeuSerPro 168
              |||       |||||||||          |||      :::   ::::::||||||
Db        652 CACGGCGCCCAGCAGCAAGCGCCAGCAGCAGCTGGCACGTCGGCGGCGGCGGTCAGCCCC 711

Qy        169 GlnLeuGlnGlnGlnIleSerPheAsnAsnAsnIleSerSerSerSerProArgThrAsn 188
                |||                                 ||||||
Db        712 GTGCTGAGC-------------------------CTGTCGTCGCGTGAGGCGGCGCCG 744

Qy        189 AsnValThrIleArgGlyThrLeuAspGlySerSer--------------------Ser 201
                ||||||:::                 ::::::
Db        745 CCGGTCACCGTGGCCGCCGCCGCCGCCGCCGCCGTCCCCGGCGGCACGGACCAGGAG 804

Qy        202 AsnMetValLeuGlySerLysTyrLeuLysAlaAlaGlnGluLeuLeuAspGluValVal 221
                :::|||:::   |||:::|||||||||||||||||||||||||||||||   |||
Db        805 AAGGTGGTGATGCGCTCCAGGTACCTGAAGGCGGCGCAGGAGCTGCTCGACGAGGCGGTG 864

Qy        222 AsnIleValGlyLysSerIleLysGlyAspAspGlnLysLysAspAsnSerMetAsnLys 241
                :::   |||   |||                 :::||||||::::::|||
Db        865 AGC---GTCAGCAAGGGCGCCGCGACGGCGGTGAAGAAGAAGGAGGACTCGGAGGGCGGC 921

Qy        242 GluSerMetProLeuAlaSerAspValAsnThrAsnSerSerGlyGlyGlyGluSerSer 261
                |||                                 ||||||   |||
Db        922 GTGTCC-------------------------GGTGGCGGCGGCGGTGCGGAGGACGGT 954

Qy        262 SerArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGlnArgGlnGluLeuGln 281
                  |||:::    |||   |||::::::||||||:::|||||||||||||||
Db        955 GGCGGCAGCAAGAGCGGCGCCGCCGCCGAGATGTCGACGGCGGAGCGGCAGGAGCTGCAG 1014

Qy        282 MetLysLysAlaLysLeuLeuAlaMetLeuGluGluValGluGlnArgTyrArgGlnTyr 301
              |||||||||:::|||||||||   ||||||:::||||||||||||||||||||||||||
Db       1015 ATGAAGAAGAGCAAGCTACTGAACATGCTCGATGAGGTGGAGCAGCGGTACAGGCAGTAC 1074

Qy        302 HisHisGlnMetGlnIleIleValLeuSerPheGluGlnValAlaGlyIleGlySerAla 321
              |||   |||||||||   :::      :::||||||      ||||||   |||||||||
Db       1075 CACCGGCAGATGCAGGGGGTGGCGGCGGCGTTCGAGGCGGCGGCGGGGGCCGGGTCGGCG 1134

Qy        322 LysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArgCysLeuLysAsp 341
              :::||||||   |||||||||       ||||||:::||||||||||||||||:::|||
Db       1135 ACGACGTACACGTCGCTGGCGCTGCGCACCATCTCGCGGCAGTTCCGGTGCCTCCGCGAC 1194

Qy        342 AlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGlu------------Glu 357
              |||||||||   ||||||:::|||   |||:::   |||||||||            :::
Db       1195 GCCATCGCGGCGCAGGTCAGGGCGGCGAGCCGGGGGCTCGGGGAGGACTGCGGCGACGAC 1254

Qy        358 GluGlyLeuGlyGly---------LysIleGluGlySerArgLeuLysPheValAspHis 374
              ||||||   ||||||                ||||||||||||:::|||:::|||||||
Db       1255 GAGGGCGGCGGTGGCGGCGGGAGGACGACGGTGGGGTCGCGGCTGCGGTTCATCGACCAC 1314

Qy        375 HisLeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGlnProAsnAla------ 392
              |||||||||||||||||:::||||||:::||||||:::          |||
Db       1315 CAGCTCCGGCAGCAGCGCGCGATGCAGCAGCTTGGCATGGTGCACGCCGCCGCCGGT 1374

Qy        393 -----------------TrpArgProGlnArgGlyLeuProGluArgAlaValSerVal 406
                              ||||||||||||||||||||||||||||||||||||||||||
Db       1375 GGCGCCGCCGGCGGCGGGTGGCGGCCGCAGCGTGGCCTCCCCGAGCGCGCCGTCTCCGTC 1434

Qy        407 LeuArgAlaTrpLeuPheGluHisPheLeuHisProTyrProLysAspSerAspLysIle 426
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db       1435 CTCCGCGCTTGGCTCTTCGAGCATTTCCTCCACCCATACCCCAAGGATTCGGACAAGGTC 1494

Qy        427 MetLeuAlaLysGlnThrGlyLeuThrArgSerGln---------------------- 438
              ||||||||||||||||||||||||||||||||
```

```
Db       1495 ATGCTCGCCAAGCAAACCGGCCTCACCAGGAGCCAGGTACGTACGTATACGTCTACGTAC 1554

Qy        439 ----------------------------------------------ValSerAsnTrpPhe 443
                                                            ||||||||||||||
Db       1555 AATACTACGTCTACGTATGTTTGAAGTACGTGTACGTATTTGTAGGTGTCGAACTGGTTC 1614

Qy        444 IleAsnAlaArgValArgLeuTrpLysProMetValGluGluMetTyrLeuGluGluVal 463
               ||||||||||||||||||||||||||||||||||||||||||||||||||    ||||||
Db       1615 ATCAACGCGAGGGTGAGGCTGTGGAAGCCAATGGTGGAGGAGATGTACGCCGAGGAGACC 1674

Qy        464 LysAsnGlnGluGlnAsnSerThrAsnThrSerGlyAspAsnLysAsnLysGluThrAsn 483
               |||   :::|||:::
Db       1675 AAGGCCAAGGAGGAG---------------------------------------------- 1689

Qy        484 IleSerAlaProAsnGluGluLysHisProIleIleThrSerSerLeuLeuGlnAspGly 503
                         ||||||:::|||           ::::::           |||
Db       1690 ------------GAGGAGGAGGAGCACGACGCGGCGGCGGCCGCCGCCGGCGACCGTGGC 1737

Qy        504 IleThrThrThrGlnAla-----------------GluIleSerThrSerThrIleSer 517
                         ||||||                                  |||
Db       1738 GGCGTGGCGGAGCAGGCTCCGAGTAAGCCGGACGACAGTGCTGGCATTGGCATGTCGTCG 1797

Qy        518 ThrSerProThrAlaGlyAlaSer------------LeuHis------------HisAla 529
               :::||||||       ||||||           :::|||           ||||||
Db       1798 TCGTCGCCGGCGGTGGCGGCGTCCAGGTCAGTTGGTGTGCACGCCGGAGACCAGCACGCT 1857

Qy        530 HisAsnPheSerPheLeuGlySerPheAsnMetAspAsn-------------------- 542
                        :::   |||             |||:::
Db       1858 CAGGCGAGCTTCTACGGCGGCGGCGGCGGCGGCGACGATCCGTTCCAGTGCAGGATCAAG 1917

Qy        543 ------ThrThrThrThrValAsp----------------------HisIleGluAsn 552
                    |||||||||    |||                      |||:::
Db       1918 AAGGCGAGGACGACGACGGCGGACGAGCCGGCGGCGGCCGCGGCGTTCCACGTCTCCGGC 1977

Qy        553 AsnAlaLys---LysGlnArgAsnAspMetHisLysPheSerProSerSer--------- 568
               |||          |||    :::   ||||||:::   :::
Db       1978 GAGGCCGCCGTGAGCCACCGGGAGCTGCTAATGAAGTTCACTGAGGCCGGCGGCGAGGGC 2037

Qy        569 ------------------IleLeuSerSerValAspMetGluAlaLysAlaArgGluSer 582
                               ::::::     |||   :::
Db       2038 GTGAGAACGGGCCACCCGCACGTCAACGACGACGACGACGACGTCCCTGGCGGCGCCGGC 2097

Qy        583 SerAsnLysGlyPheThrAsnProLeuMetAlaAlaTyrAlaMetGlyAspPheGlyArg 602
                         ||||||            ||||||                ||||||
Db       2098 GGGTACTCGCTGTTCACG-----------GCGGCGCAGTACGGGCATCAGTTCGGC--- 2142

Qy        603 PheAspProHisAspGlnGlnMetThrAlaAsnPheHisGlyAsnAsnGlyValSerLeu 622
                  |||                              |||      ||||||||||||
Db       2143 ---TCGGACCATTTCGCGTTCGCCGGGCACGGCGGCGGAGGCGGCGGCGGCGTGTCGCTC 2199

Qy        623 ThrLeuGlyLeuProProSerGluAsnLeuAlaMetProValSerGlnGlnAsnTyrLeu 642
               ||||||||||||||         :::        |||   |||         :::|||
Db       2200 ACGCTCGGCCTCCCGCACGGCGCCGAC---CAGACGCCGGCGTCG---------TTTCTC 2247

Qy        643 SerAsnAspLeuGlySerArgSerGluMetGlySerHisTyrAsnArgMetGlyTyrGlu 662
                                        |||:::                ||||||:::
Db       2248 ATAGGCGCCGGCGCCGGCAGCGACGGCGGCGGCGCGCCGGTGACGACGGCCGGCTACGAC 2307

Qy        663 AsnIleAspPheGlnSerGlyAsnLysArgPheProThrGlnLeuLeuProAspPheVal 682
                  ::::::   ||||||      |||            ||||||:::   |||||||||
Db       2308 ---ATGAACATGCAGAGC---ACCAAGTCGTTGGCTGCTCAGCTCATGAGAGACTTCGTG 2361
```

RESULT 15
AK070465
LOCUS AK070465 3030 bp mRNA linear PLN 24-JUL-2003
DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:J023054P21, full insert sequence.
ACCESSION AK070465
VERSION AK070465.1 GI:32980489
KEYWORDS FLI_CDNA; CAP trapper.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1
AUTHORS The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team:, Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group:, Otomo,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y., Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M., Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J., Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J., Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN:, Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y., Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Yoshino,M. and Hayashizaki,Y.
TITLE Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice
JOURNAL Science 301 (5631), 376-379 (2003)
MEDLINE 22752273
PUBMED 12869764
REFERENCE 2 (bases 1 to 3030)
AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K., Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K., Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M., Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M., Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M., Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A., Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M., Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K., Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H., Osato,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K., Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S., Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W., Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and Yoshimura,A.
TITLE Direct Submission
JOURNAL Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)
COMMENT This clone is one of the 28K full-length cDNA clones from japonica rice.
URL : http://cdna01.dna.affrc.go.jp/cDNA/
NIAS Rice Full-Length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and Yamamoto,M.
FAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y., Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M., Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J., Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S., Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S., Yoshimura,A., Matsubara,K. and Murakami,K.
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J.,

Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N., Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K., Yasunishi,A. and Hayashizaki,Y.

FEATURES Location/Qualifiers
source 1. .3030
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J023054P21"

ORIGIN

Alignment Scores:

Pred. No.:	4.72e-59	Length:	3030
Score:	945.00	Matches:	252
Percent Similarity:	50.79%	Conservative:	100
Best Local Similarity:	36.36%	Mismatches:	215
Query Match:	26.52%	Indels:	126
DB:	8	Gaps:	23

US-10-624-201A-2 (1-688) x AK070465 (1-3030)

```
Qy        33 LeuTyrLeuMetAsnProAsnAsnTyrMetGlnGlyTyrThrThr---SerAspThrGln 51
             |||||||||   ||||||::::::           ||||||     |||
Db       897 CTCTATCTGAACAATCCATCTTCT------GGACCTTATACGGAGTTCAGCGGCATTCTG 950

Qy        52 GlnGlnGlnGlnLeuLeuPheLeuAsnSerSerProAlaAlaSerAsnAlaLeuCysHis 71
             |||   ||||||                 |||
Db       951 CAGCCTCAGCAGAATTGCATGGAGATGCCTGGCCCTGGG-------------------- 989

Qy        72 AlaAsnIleGlnHisAlaProLeuGlnGlnGlnHisPheValGlyValProLeuProAla 91
                        ||||||              |||                    |||:::
Db       990 ------------CATGCGTCAGCAATGTCACAAGAC-----------------CCCTCT 1019

Qy        92 ValSerLeuHisAspGlnIleAsnHisHisGlyLeuLeuGlnArgMetTrpAsnAsnGln 111
                  |||  ::::::   |||       :::     :::   :::
Db      1020 TCAAGGGAGTCCGACATGCTTTCTTCGCACCAGGGTCAACGCTCTTTCAGCCATGTCAAA 1079

Qy       112 AspGlnSerGlnGlnValIleValProSerSerThrGlyValSerAlaThrSerCysGly 131
             |||          ::::::::::::           |||          :::
Db      1080 GATATGAAAAATGAGATGTTGATGCATATGATGGATGGAGCACAAGGTAGTGGCTCTGAG 1139

Qy       132 GlyIleThrThrAsp------LeuAlaSerGlnLeuAlaPhe----------------- 143
                |||      |||            |||||||||   |||
Db      1140 CTCATCCATGATGATGCCCACACTGGTTCACAGCTTGAGTTTGGTGTTCTGAATAATCAC 1199

Qy       144 ---GlnArgProIleProThrProGlnHisArgGlnGlnGlnGlnGlnGlyGlyLeu 162
                   :::|||:::                        |||      ||||||
Db      1200 AACTCATCAAGTGTTCCATCA--------------------ATGCAAAGCCAAGGGCTA 1238

Qy       163 SerLeuSerLeuSer-------------------------------------------Pro 168
             ||||||||||||:::                                          |||
Db      1239 TCTCTGAGCTTGAACACACAAATCATGGCGCCTTCCTTACCGTACTGGTCTATAAAGCCA 1298

Qy       169 GlnLeuGlnGlnGlnIleSerPheAsnAsnAsnIleSerSerSerSerProArgThrAsn 188
                :::            |||:::::::::|||:::              |||
Db      1299 GACATGTTAACGCCGCAGTCTTACCATGACAACCTTAGAGGGGAAGACATGCGGATGAAG 1358

Qy       189 AsnValThrIleArgGlyThrLeuAspGlySerSerSerAsnMetValLeuGlySerLys 208
             |||                |||          :::|||     :::     |||:::
Db      1359 AAT-----------------TTGCAGTCCGAAGCTTCACGCGCAATCCGGAACTCAAGG 1400

Qy       209 TyrLeuLysAlaAlaGlnGluLeuLeuAspGluValValAsnIleValGlyLysSerIle 228
             |||||||||||||||||||||||||||||||||||:::   |||   |||||||||
Db      1401 TATCTCAAGGCAGCACAAGAACTGCTTGATGAGGTCGTGAGT---GTTTGGAAGAGCATA 1457

Qy       229 LysGlyAspAspGlnLys-----------------LysAspAsnSerMetAsnLysGlu 242
             |||          ||||||                |||   :::           :::
```

```
Db      1458 AAACAAAAGGCTCAGAAAGAGAAGGTTGAATCAGGAAAAGCAGATGGCAAAGAGACTGAT 1517

Qy       243 SerMetProLeuAlaSerAspValAsnThrAsnSerSerGlyGlyGlyGluSerSerSer 262
                |||   :::       |||::::::|||                |||   ::::::
Db      1518 GGAGGGCCAAAAAGTGAGGGTGTATCTTCTAATCCACAAGAATCTGGCGCTAATGCA--- 1574

Qy       263 ArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGlnArgGlnGluLeuGlnMet 282
                                 |||   ||||||:::||||||:::::|||||||||||||
Db      1575 -----------------GCACCTGAGCTTTCCACTGCTGAGAAGCAAGAGCTTCAGAAC 1616

Qy       283 LysLysAlaLysLeuLeuAlaMetLeuGluGluValGluGlnArgTyrArgGlnTyrHis 302
             |||   |||||||||:::|||||||||:::||||||:::::::::|||:::   |||:::
Db      1617 AAAATGGCAAAACTGATGGCAATGTTGGATGAGGTGGACCGGAAATACAAGCATTATTAC 1676

Qy       303 HisGlnMetGlnIleIleValLeuSerPheGluGlnValAlaGlyIleGlySerAlaLys 322
             ||||||||||||   :::|||   ||||||:::   |||||||||   ||||||||||
Db      1677 CACCAAATGCAAACCGTGGTTTCATCGTTTGATGTGGTAGCTGGACCAGGATCTGCGAAG 1736

Qy       323 SerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArgCysLeuLysAspAla 342
                 ||||||   :::||||||      ||||||:::   ||||||||||||||||||||
Db      1737 CCTTACACAGCAGTTGCTCTTCAGACAATCTCACGGCACTTTCGGTGCCTGAAGGATGCT 1796

Qy       343 IleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGluGluGluGlyLeuGlyGly 362
             |||   :::|||:::              |||   ||||||||||||||||          |||
Db      1797 ATCAATGATCAGATCAATGTTATCAGGAAAAAGCTTGGCGAGGAAGAGAACTCATCTGGC 1856

Qy       363 LysIleGluGly------SerArgLeuLysPheValAspHisHisLeuArgGlnGlnArg 380
             |||   ||||||      :::||||||:::::::::|||        ||||||||||||||
Db      1857 AAG---GAGGGAAAATTAACCCGCCTCCGGTACATTGATCAGCAGCTAAGGCAACAACGT 1913

Qy       381 AlaLeuGlnGlnIleGlyMetMetGlnProAsnAlaTrpArgProGlnArgGlyLeuPro 400
             |||   ||||||   ||||||:::       ||||||||||||||||||||||||||||||
Db      1914 GCTTTCCAACAGTACGGTATGATTCCACAAAACGCCTGGAGACCACAGAGGGGACTGCCT 1973

Qy       401 GluArgAlaValSerValLeuArgAlaTrpLeuPheGluHisPheLeuHisProTyrPro 420
             |||   :::|||::::::|||||||||||||||||||||||||||||||||||||||||
Db      1974 GAAAACTCTGTTACAATTCTTCGCGCTTGGCTGTTTGAACACTTCCTCCATCCGTACCCA 2033

Qy       421 LysAspSerAspLysIleMetLeuAlaLysGlnThrGlyLeuThrArgSerGlnValSer 440
             |||||||||:::|||:::|||||||||:::|||||||||||||||||||||||:::|||
Db      2034 AAAGATTCCGAGAAGTTAATGCTTGCTAGGCAGACTGGCCTGACTAGGAGTCAGATTTCG 2093

Qy       441 AsnTrpPheIleAsnAlaArgValArgLeuTrpLysProMetValGluGluMetTyrLeu 460
             |||||||||||||||||||||||||||||||||||||||:::|||:::||||||
Db      2094 AATTGGTTCATAAATGCCCGTGTCCGCCTTTGGAAACCAATGATTGAAGACATGTACAAA 2153

Qy       461 GluGluValLysAsnGlnGluGlnAsnSerThrAsnThrSerGlyAspAsnLysAsnLys 480
             ||||||:::   :::   ||||||:::|||   |||:::|||   ||||||      :::
Db      2154 GAAGAGATTGGTGATTTAGAGCAAGACTCC---AACTCTTCCTCTGACAATGCACCAAGA 2210

Qy       481 GluThrAsnIleSerAlaProAsnGluGluLysHisProIleIleThrSerSerLeuLeu 500
                :::      |||   :::|||:::|||       :::      |||
Db      2211 AGTAAGGATAAAATGGCATCTTCAGAAGACAAGGAGGATCTCAAAAACTCCAGGGCACGG 2270

Qy       501 GlnAspGlyIleThrThrThrGlnAlaGluIleSerThrSerThrIleSerThrSerPro 520
                         |||:::|||      |||   :::::::|||
Db      2271 ATT------TGCGAAACAAGCCAATTGAGCGAGTCGAGAACCAGCATTGGAGCAATGAAT 2324

Qy       521 ThrAlaGlyAlaSerLeuHisHisAlaHis------------AsnPheSerPheLeuGly 536
                ||||||   :::          :::             :::   ||||||:::
Db      2325 GTTGGCGGAGCGCCTGTTGGCTTCCAGAATGAGCCCAATCCTGATGACAGTTTCATGAAC 2384

Qy       537 SerPheAsnMetAspAsnThrThrThrThrValAspHis------IleGluAsnAsnAla 554
                         |||          |||||          :::   :::|||
Db      2385 CTGATGCTGAAGGACCAGCGATCGAACGAAGTGGATGGCGGCCTCCTTCTCCACAACACT 2444

Qy       555 LysLysGlnArgAsnAspMetHisLysPheSerProSerSerIleLeuSerSerValAsp 574
                ||||   :::|||
Db      2445 GTCGCGCAGCACTCAGAC---------------------------------------- 2462

Qy       575 MetGluAlaLysAlaArgGluSerSerAsnLysGlyPheThrAsnProLeuMetAlaAla 594
                                                      |||          |||
Db      2463 ------------------------------------GAGAACGCCCGGTTCATGGCC 2483
```

```
Qy        595 TyrAlaMetGlyAspPheGlyArgPheAspProHisAspGlnGlnMetThrAlaAsnPhe 614
              |||   :::   :::   ||||||:::
Db       2484 TATCACTTGGCAGAGCTTGGGAGATAC------------------------------- 2510

Qy        615 HisGlyAsnAsnGlyValSerLeuThrLeuGlyLeuProProSerGluAsnLeuAlaMet 634
                 ||||||      ||||||||||||||||||||      |||   :::      :::
Db       2511 ---GGCAATGGCAATGTGTCGCTGACACTGGGCTTGCAGCACTCCAGCAGCAACCTCGTT 2567

Qy        635 ProValSerGlnGlnAsnTyr------LeuSerAsnAsp------LeuGlySerArgSer 650
              |||   :::|||      :::      ::::::   |||           :::   :::
Db       2568 CCTAATGCTCAGCCAGGTTTCCCTGGTGTTAATGAGGATGACATCTACAATGCTACTGCT 2627

Qy        651 GluMetGlySerHisTyrAsnArgMetGlyTyrGluAsnIleAspPheGlnSerGlyAsn 670
                :::|||                        |||::::::::::::
Db       2628 CCTCTTGGTGTCACCGTCGCATCTTCGGACTATGATTCGATGAATCAGATGGATCAACGG 2687

Qy        671 LysArgPhePro---ThrGlnLeuLeuProAspPheVal 682
              :::||||||      :::   ||||||   |||||||||
Db       2688 CAACGGTTTGAGCATTCACCTCTTCTGCATGATTTTGTG 2726
```

Search completed: July 29, 2005, 17:21:49
Job time : 8239 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 29, 2005, 08:33:52 ; Search time 949 Seconds
(without alignments)
4291.657 Million cell updates/sec

Title: US-10-624-201A-2
Perfect score: 3563
Sequence: 1 MYYQGTSDNTNIQADHQQRH.....GNKRFPTQLLPDFVTGNLGT 688

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10624201/runat_27072005_122956_10897/app_query.fasta_1.839
-DB=N_Geneseq_16Dec04 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10624201_@CGN_1_1_644_@runat_27072005_122956_10897 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1150.5	32.3	2043	6	ABZ13532	Abz13532 Arabidops
2	1150.5	32.3	2043	9	ADB23149	Adb23149 Environme
3	1148.5	32.2	2385	10	ADD30303	Add30303 Plant yie
4	1148.5	32.2	2385	12	ADI44208	Adi44208 Plant tra
5	926.5	26.0	1929	12	ADQ37104	Adq37104 Cell prol
6	826	23.2	1971	12	ADO62706	Ado62706 Transcrip
7	824.5	23.1	1617	12	ADO63042	Ado63042 Transcrip
8	824	23.1	1884	6	ABZ13100	Abz13100 Arabidops
9	824	23.1	1983	4	AAD06494	Aad06494 Arabidops
10	824	23.1	1983	10	ADC46618	Adc46618 Thalecres
11	824	23.1	1983	10	ADD30786	Add30786 Plant yie
12	824	23.1	1983	12	ADI44018	Adi44018 Plant tra
13	824	23.1	1983	12	ADO02224	Ado02224 Thalecres
14	820	23.0	2025	3	AAC39138	Aac39138 Arabidops
15	816	22.9	2458	12	ADO62707	Ado62707 Transcrip
16	803	22.5	1824	3	AAC36745	Aac36745 Arabidops
17	802	22.5	1449	3	AAC43406	Aac43406 Arabidops
18	802	22.5	1449	6	ADG88236	Adg88236 A. thalia
19	802	22.5	1575	4	AAD06493	Aad06493 Arabidops
20	802	22.5	1575	12	ADO61968	Ado61968 Transcrip
21	802	22.5	1599	6	ABZ14349	Abz14349 Arabidops
22	802	22.5	2031	3	AAC36460	Aac36460 Arabidops
23	797.5	22.4	2545	4	AAD06479	Aad06479 Arabidops
24	797.5	22.4	2545	10	ADE37296	Ade37296 Plant yie
25	797.5	22.4	2545	12	ADI43866	Adi43866 Plant tra
26	797.5	22.4	2545	12	ADO01854	Ado01854 Thalecres
27	786	22.1	1485	3	AAC50215	Aac50215 Arabidops
28	732.5	20.6	1836	6	ABZ12846	Abz12846 Arabidops
29	732.5	20.6	1836	6	ADG88038	Adg88038 A. thalia
30	725	20.3	2405	4	AAD06449	Aad06449 Arabidops
31	725	20.3	2405	10	ADE37116	Ade37116 Plant yie
32	725	20.3	2405	12	ADI43868	Adi43868 Plant tra
33	725	20.3	2405	12	ADO01856	Ado01856 Thalecres
34	701.5	19.7	1296	12	ADO61794	Ado61794 Transcrip
35	678	19.0	2034	3	AAC50226	Aac50226 Arabidops
36	649.5	18.2	1905	3	AAC39451	Aac39451 Arabidops
37	646.5	18.1	2442	10	AAL55526	Aal55526 qSH-1 gen
38	646.5	18.1	2450	10	AAL55527	Aal55527 qSH-1 gen
39	639	17.9	894	12	ADQ37116	Adq37116 Cell prol
40	636	17.9	706	6	ABK82114	Abk82114 DNA encod
41	616.5	17.3	2131	6	ABL56175	Abl56175 Hordeum v
42	607	17.0	2141	6	ABL56177	Abl56177 Hordeum v
43	605	17.0	626	5	AAH87740	Aah87740 Peppermin
44	586.5	16.5	1239	3	AAC55934	Aac55934 Eucalyptu
45	493	13.8	529	3	AAC55946	Aac55946 Eucalyptu

ALIGNMENTS

RESULT 1
ABZ13532
ID ABZ13532 standard; DNA; 2043 BP.
XX
AC ABZ13532;
XX
DT 21-JAN-2003 (first entry)
XX
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 1337.
XX
KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX
OS Arabidopsis thaliana.
XX
PN WO200216655-A2.
XX
PD 28-FEB-2002.
XX
PF 24-AUG-2001; 2001WO-US026685.
XX
PR 24-AUG-2000; 2000US-0227866P.
PR 26-JAN-2001; 2001US-0264647P.
PR 22-JUN-2001; 2001US-0300111P.
XX
PA (SCRI) SCRIPPS RES INST.
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Harper JF, Kreps J, Wang X, Zhu T;
XX

DR WPI; 2002-304127/34.
XX
PT Identifying a stress condition to which a plant cell has been exposed and
PT producing plants with increased tolerance to these abiotic stresses.
XX
PS Claim 144; SEQ ID NO 1337; 577pp + Sequence Listing; English.
XX
CC The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising: (a) contacting nucleic acid
CC representative of expressed polynucleotides in the plant cell with an
CC array or probes representative of the plant cell genome; and (b)
CC detecting a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants
CC with increased tolerance to abiotic stress. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
CC in methods of the invention. Note: The sequence data for this patent is
CC not represented in the printed specification but is based on sequence
CC information supplied to Derwent by the European Patent Office
XX
SQ Sequence 2043 BP; 627 A; 507 C; 507 G; 402 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.32e-85	Length:	2043
Score:	1150.50	Matches:	306
Percent Similarity:	51.81%	Conservative:	81
Best Local Similarity:	40.96%	Mismatches:	201
Query Match:	32.29%	Indels:	159
DB:	6	Gaps:	28

US-10-624-201A-2 (1-688) x ABZ13532 (1-2043)

```
Qy        22 HisGlyAsn-----------------SerAsnAsnAsnAsnIleGlnThrLeuTyrLeu 35
             |||||||||                      ::::::      :::||||||||||   |||
Db        16 CACGGAAACCCACCGGAGATCTCTGCCGGATCCGACGGTGGTCTTCAAACGTTGATCCTC 75

Qy        36 MetAsnProAsnAsnTyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGlnGlnGln 55
             |||||||||      |||:::|||   ||||||      |||
Db        76 ATGAATCCAACTACTTACGTTCAG---TACACCCAACAAGACAACGACTCGAACAACAAC 132

Qy        56 LeuLeuPheLeuAsnSerSerProAlaAlaSerAsnAlaLeuCysHisAlaAsnIle--- 74
                    |||:::|||      ::::::|||         :::   |||
Db       133 -----------AACAACAGCAACAATAGCAACAACAACAACACAAACACAAACACAAAC 180

Qy        75 ------------------------------GlnHisAlaProLeu------GlnGlnGln 82
                                           |||||||||                |||
Db       181 AACAACAACAGTAGTTTCGTTTTCCTCGATTCCCACGCGCCGCAGCCAAACGCGAGCCAG 240

Qy        83 HisPheValGlyValProLeu--------------ProAlaValSerLeuHisAspGln 97
                |||||||||:::||||||                :::::::::      |||
Db       241 CAGTTCGTCGGAATACCACTCTCAGGTCACGAAGCTGCTTCCATTACAGCCGCCGACAAC 300

Qy        98 IleAsn---HisHisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnSerGlnGln 116
             |||:::      ||||||          |||:::      :::         |||   |||
Db       301 ATCTCCGTACTTCACGGTTATCCTCCGCGCGTGCAGTACAGTCTCTACGGTAGCCACCAA 360

Qy       117 ValIleValProSerSerThrGlyValSerAlaThrSerCysGlyGlyIleThrThrAsp 136
             |||
Db       361 GTG---------------------------------------------------------- 363

Qy       137 LeuAlaSerGlnLeuAlaPheGlnArgProIleProThrProGlnHisArgGln------ 154
                                                 |||   |||:::|||
Db       364 ----------------------------------GATCCCACTCACCAGCAAGCCGCG 387

Qy       155 ------GlnGlnGlnGlnGlnGlyGlyLeuSerLeuSerLeuSerProGlnLeuGlnGln 172
                    :::   |||  ||||||||||||:::||||||   |||   ||||||
Db       388 TGTGAGACGCCACGCGCGCAGCAAGGCCTCTCTTTAACCCTCTCGTCTCAACAGCAGCAG 447

Qy       173 GlnIleSerPheAsnAsn--------------AsnIleSerSerSerSerProArgThr 187
             |||      :::                  ::::::       |||
Db       448 CAACAGCAACATCATCAACAACACCAGCCTATTCACGTCGGATTCGGGTCCGGACATGGA 507

Qy       188 AsnAsnValThrIleArgGlyThrLeuAspGlySerSer--------------SerAsn 202
```

```
                ::::::   :::                 ||||||                 :::|||
Db       508 GAAGATATCCGGGTCGGGTCTGGCTCTACAGGATCGGGGGTAACAAACGGTATAGCTAAT 567

Qy       203 MetValLeuGlySerLysTyrLeuLysAlaAlaGlnGluLeuLeuAspGluValValAsn 222
             :::|||      ||||||||||||||||||||||||||||||||||||||||||||||||
Db       568 CTTGTT---AGCTCCAAGTACTTGAAGGCAGCACAAGAGCTTCTTGACGAAGTAGTCAAC 624

Qy       223 IleValGlyLysSerIleLysGlyAspAspGln------------LysLysAspAsnSer 238
                           :::                |||          ||||||   :::
Db       625 GCTGATTCCGATGACATGAACGCTAAATCCCAACTATTCTCATCGAAAAAGGGTAGTTGC 684

Qy       239 MetAsnLysGluSerMetProLeuAlaSer---AspValAsnThrAsnSerSerGlyGly 257
               |||   :::   :::       :::|||                         |||||||||
Db       685 GGAAATGATAAACCTGTCGGAGAATCATCGGCCGGCGCTGGAGGAGAAGGTTCCGGTGGC 744

Qy       258 GlyGluSerSerSerArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGlnArg 277
             |||      ::::::   ::::::            |||||||||   ||||||:::|||
Db       745 GGAGCAGAAGCAGCCGGGAAACGT---------CCGGTGGAGCTAGGCACGGCAGAGAGA 795

Qy       278 GlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeuGluGluValGluGlnArg 297
             ||||||:::|||||||||||||||||||      ||||||   ||||||||||||||||
Db       796 CAAGAAATACAGATGAAGAAAGCAAAACTTAGTAACATGCTTCATGAGGTGGAGCAGAGA 855

Qy       298 TyrArgGlnTyrHisHisGlnMetGlnIleIleValLeuSerPheGluGlnValAlaGly 317
             ||||||||||||||   |||||||||:::::::::   ||||||||||||   ||||||
Db       856 TATAGACAGTACCACCAGCAGATGCAGATGGTGATCTCTTCGTTCGAGCAAGCGGCAGGG 915

Qy       318 IleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArg 337
             ||||||||||||||||||||||||   |||||||||      |||||||:::|||||||||
Db       916 ATAGGATCAGCGAAGTCATACACGTCGCTAGCATTGAAAACCATATCAAGACAGTTCCGT 975

Qy       338 CysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGluGlu 357
             |||||||||:::|||||||||   |||:::||||||   :::|||||||||||||||||
Db       976 TGCTTGAAAGAGGCGATCGCTGGTCAGATAAAAGCGGCCAACAAGAGTCTTGGGGAGGAA 1035

Qy       358 GluGlyLeuGly------GlyLysIleGluGlySerArgLeuLysPheValAspHisHis 375
             :::   :::         |||:::   ||||||||||||||||||||||||||||||||
Db      1036 GATTCAGTGTCTGGTGTTGGGAGGTTTGAGGGGTCGAGGCTCAAGTTCGTGGACCACCAC 1095

Qy       376 LeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGln---------ProAsnAla 392
             ||||||||||||||||||||||||||::||||||:::|||           ||||||
Db      1096 TTGAGACAGCAAAGAGCTCTTCAACAACTGGGAATGATTCAACATCCTTCCAATAATGCT 1155

Qy       393 TrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeuPhe 412
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1156 TGGAGACCTCAACGTGGTCTCCCAGAACGAGCCGTCTCAGTTCTCCGTGCTTGGCTCTTC 1215

Qy       413 GluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGlnThr 432
             ||||||||||||||||||||||||||||||||||||   ||||||||||||||||||
Db      1216 GAACACTTTCTTCATCCATACCCTAAGGATTCGGACAAGCACATGCTAGCTAAGCAAACA 1275

Qy       433 GlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLys 452
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1276 GGACTCACTCGTAGCCAGGTGTCGAACTGGTTTATAAACGCGAGAGTTCGGTTATGGAAA 1335

Qy       453 ProMetValGluGluMetTyrLeuGluGluValLysAsnGlnGluGlnAsnSerThrAsn 472
             ||||||||||||||||||||:::||||||:::|||   |||   :::|||
Db      1336 CCAATGGTGGAGGAGATGTACATGGAGGAAATGAAGGAGCAGGCAAAGAACATG------ 1389

Qy       473 ThrSerGlyAspAsnLysAsnLysGluThrAsnIleSerAlaProAsnGluGluLysHis 492
                                    ::::::|||   :::          ||||||:::
Db      1390 ------------GGATCCATGGAAAAGACTCCTTTGGATCAAAGCAACGAAGAT------ 1431

Qy       493 ProIleIleThrSerSerLeuLeuGlnAspGlyIleThrThrThrGlnAlaGluIleSer 512
                                                                |||   |||
Db      1432 ---------------------------------------------TCTGCTTCAAAGTCA 1446

Qy       513 ThrSerThrIleSerThrSerProThrAlaGlyAlaSerLeuHis------HisAlaHis 530
             ||||||             ||||||   |||      :::   |||           |||
Db      1447 ACAAGTAACCAAGAAAAGAGCCCAATGGCGGACACTAATTACCATATGAATCCCAATCAC 1506

Qy       531 AsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrThrValAspHisIle 550
             |||            |||         |||      :::         :::
```

```
Db       1507 AACGGTGACCTAGAAGGCGTCACTGGAATGCAAGGAAGCCCCAAGAGACTAAGAACCAGC 1566

Qy        551 GluAsnAsnAlaLysLysGlnArgAsnAspMetHisLysPheSerProSerSerIleLeu 570
Db       1567 GACGAGACAATGATGCAGCCAATAAAT------GCGGATTTCAGCTCCAACGAGAAGCTC 1620

Qy        571 SerSerValAspMetGluAlaLysAlaArgGluSerSerAsnLysGlyPheThrAsnPro 590
Db       1621 ACGATGAAGATTCTAGAAGAACGGCAAGGGATAAGATCAGATGGTGGCTAC------CCT 1674

Qy        591 LeuMet---------AlaAlaTyrAlaMetGlyAspPheGlyArgPheAsp---ProHis 606
Db       1675 TTCATGGGTAATTTCGGGCAATACCAAATGGATGAGATGTCAAGATTTGATGTAGTCTCA 1734

Qy        607 AspGlnGlnMetThrAlaAsnPheHis------GlyAsnAsnGlyValSerLeuThrLeu 624
Db       1735 GACCAGGAGCTCATGGCGCAAAGGTACTCAGGAAACAACAATGGCGTGTCCCTCACGTTA 1794

Qy        625 GlyLeuProProSerGluAsnLeuAlaMetProValSerGln---------GlnAsnTyr 641
Db       1795 GGTTTACCTCATTGTGATAGCTTGTCGTCCACGCACCATCAGGGTTTCATGCAGACCCAC 1854

Qy        642 LeuSerAsnAspLeuGlySerArgSerGluMetGly---------------------- 653
Db       1855 CATGGGATTCCTATAGGGAGAAGAGTGAAAATAGGAGAAACAGAGGAATATGGACCCGCC 1914

Qy        654 ----------------------------------------------------SerHis 655
Db       1915 ACCATCAATGGTGGTAGCTCGACCACAACCGCACATTCATCAGCGGCAGCTGCCGCGGCT 1974

Qy        656 TyrAsnArgMetGlyTyrGluAsnIleAspPheGlnSerGlyAsnLysArgProProThr 675
Db       1975 TACAATGGGATGAACATACAGAAC-----------------CAGAAGAGATATGTGGCT 2016

Qy        676 GlnLeuLeuProAspPheVal 682
Db       2017 CAGTTATTGCCCGACTTCGTT 2037
```

RESULT 2
ADB23149
ID ADB23149 standard; DNA; 2043 BP.
XX
AC ADB23149;
XX
DT 20-NOV-2003 (first entry)
XX
DE Environmental stress-responsive promoter-related DNA, SEQ ID 153.
XX
KW Plant; environmental stress; promoter; gene; ds.
XX
OS Arabidopsis thaliana.
XX
PN WO2003044190-A1.
XX
PD 30-MAY-2003.
XX
PF 15-NOV-2002; 2002WO-JP011955.
XX
PR 19-NOV-2001; 2001JP-00353038.
PR 29-JAN-2002; 2002JP-00020329.
XX
PA (RIKE) RIKEN KK.
XX
PI Shinozaki K, Seki M, Fujita M;
XX
DR WPI; 2003-468771/44.
DR P-PSDB; ADB23150.
XX
PT Environmental stress-responsive promoters, useful in constructing and
PT breeding stress-resistant plants for applications in agriculture.
XX
PS Claim 10; Page 470-475; 520pp; Japanese.
XX

CC The present invention relates to novel environmental stress-responsive
CC promoters (ADB23163-ADB23252) from Arabidopsis thaliana. The promoters
CC are useful in constructing and breeding stress-resistant plants for
CC applications in agriculture and horticulture. The present sequence was
CC used to illustrate the invention.
XX
SQ Sequence 2043 BP; 627 A; 507 C; 507 G; 402 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.32e-85	Length:	2043
Score:	1150.50	Matches:	306
Percent Similarity:	51.81%	Conservative:	81
Best Local Similarity:	40.96%	Mismatches:	201
Query Match:	32.29%	Indels:	159
DB:	9	Gaps:	28

US-10-624-201A-2 (1-688) x ADB23149 (1-2043)

```
Qy         22 HisGlyAsn-----------------SerAsnAsnAsnAsnIleGlnThrLeuTyrLeu 35
Db         16 CACGGAAACCCACCGGAGATCTCTGCCGGATCCGACGGTGGTCTTCAAACGTTGATCCTC 75

Qy         36 MetAsnProAsnAsnTyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGlnGlnGln 55
Db         76 ATGAATCCAACTACTTACGTTCAG---TACACCCAACAAGACAACGACTCGAACAACAAC 132

Qy         56 LeuLeuPheLeuAsnSerSerProAlaAlaSerAsnAlaLeuCysHisAlaAsnIle--- 74
Db        133 -----------AACAACAGCAACAATAGCAACAACAACAACACAAACACAAACACAAAC 180

Qy         75 ----------------------------GlnHisAlaProLeu------GlnGlnGln 82
Db        181 AACAACAACAGTAGTTTCGTTTTCCTCGATTCCCACGCGCCGCAGCCAAACGCGAGCCAG 240

Qy         83 HisPheValGlyValProLeu--------------ProAlaValSerLeuHisAspGln 97
Db        241 CAGTTCGTCGGAATACCACTCTCAGGTCACGAAGCTGCTTCCATTACAGCCGCCGACAAC 300

Qy         98 IleAsn---HisHisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnSerGlnGln 116
Db        301 ATCTCCGTACTTCACGGTTATCCTCCGCGCGTGCAGTACAGTCTCTACGGTAGCCACCAA 360

Qy        117 ValIleValProSerSerThrGlyValSerAlaThrSerCysGlyGlyIleThrThrAsp 136
Db        361 GTG---------------------------------------------------------- 363

Qy        137 LeuAlaSerGlnLeuAlaPheGlnArgProIleProThrProGlnHisArgGln------ 154
Db        364 ----------------------------------GATCCCACTCACCAGCAAGCCGCG 387

Qy        155 ------GlnGlnGlnGlnGlnGlyGlyLeuSerLeuSerLeuSerProGlnLeuGlnGln 172
Db        388 TGTGAGACGCCACGCGCGCAGCAAGGCCTCTCTTTAACCCTCTCGTCTCAACAGCAGCAG 447

Qy        173 GlnIleSerPheAsnAsn--------------AsnIleSerSerSerSerProArgThr 187
Db        448 CAACAGCAACATCATCAACAACACCAGCCTATTCACGTCGGATTCGGGTCCGGACATGGA 507

Qy        188 AsnAsnValThrIleArgGlyThrLeuAspGlySerSer--------------SerAsn 202
Db        508 GAAGATATCCGGGTCGGGTCTGGCTCTACAGGATCGGGGGTAACAAACGGTATAGCTAAT 567

Qy        203 MetValLeuGlySerLysTyrLeuLysAlaAlaGlnGluLeuLeuAspGluValValAsn 222
Db        568 CTTGTT---AGCTCCAAGTACTTGAAGGCAGCACAAGAGCTTCTTGACGAAGTAGTCAAC 624

Qy        223 IleValGlyLysSerIleLysGlyAspAspGln-----------LysLysAspAsnSer 238
Db        625 GCTGATTCCGATGACATGAACGCTAAATCCCAACTATTCTCATCGAAAAAGGGTAGTTGC 684

Qy        239 MetAsnLysGluSerMetProLeuAlaSer---AspValAsnThrAsnSerSerGlyGly 257
Db        685 GGAAATGATAAACCTGTCGGAGAATCATCGGCCGGCGCTGGAGGAGAAGGTTCCGGTGGC 744
```

```
Qy        258 GlyGluSerSerSerArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGlnArg 277
Db        745 GGAGCAGAAGCAGCCGGGAAACGT---------CCGGTGGAGCTAGGCACGGCAGAGAGA 795

Qy        278 GlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeuGluGluValGluGlnArg 297
Db        796 CAAGAAATACAGATGAAGAAAGCAAAACTTAGTAACATGCTTCATGAGGTGGAGCAGAGA 855

Qy        298 TyrArgGlnTyrHisHisGlnMetGlnIleIleValLeuSerPheGluGlnValAlaGly 317
Db        856 TATAGACAGTACCACCAGCAGATGCAGATGGTGATCTCTTCGTTCGAGCAAGCGGCAGGG 915

Qy        318 IleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArg 337
Db        916 ATAGGATCAGCGAAGTCATACACGTCGCTAGCATTGAAAACCATATCAAGACAGTTCCGT 975

Qy        338 CysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGluGlu 357
Db        976 TGCTTGAAAGAGGCGATCGCTGGTCAGATAAAAGCGGCCAACAAGAGTCTTGGGGAGGAA 1035

Qy        358 GluGlyLeuGly------GlyLysIleGluGlySerArgLeuLysPheValAspHisHis 375
Db       1036 GATTCAGTGTCTGGTGTTGGGAGGTTTGAGGGGTCGAGGCTCAAGTTCGTGGACCACCAC 1095

Qy        376 LeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGln---------ProAsnAla 392
Db       1096 TTGAGACAGCAAAGAGCTCTTCAACAACTGGGAATGATTCAACATCCTTCCAATAATGCT 1155

Qy        393 TrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeuPhe 412
Db       1156 TGGAGACCTCAACGTGGTCTCCCAGAACGAGCCGTCTCAGTTCTCCGTGCTTGGCTCTTC 1215

Qy        413 GluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGlnThr 432
Db       1216 GAACACTTTCTTCATCCATACCCTAAGGATTCGGACAAGCACATGCTAGCTAAGCAAACA 1275

Qy        433 GlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLys 452
Db       1276 GGACTCACTCGTAGCCAGGTGTCGAACTGGTTTATAAACGCGAGAGTTCGGTTATGGAAA 1335

Qy        453 ProMetValGluGluMetTyrLeuGluGluValLysAsnGlnGluGlnAsnSerThrAsn 472
Db       1336 CCAATGGTGGAGGAGATGTACATGGAGGAAATGAAGGAGCAGGCAAAGAACATG------ 1389

Qy        473 ThrSerGlyAspAsnLysAsnLysGluThrAsnIleSerAlaProAsnGluGluLysHis 492
Db       1390 -----------GGATCCATGGAAAAGACTCCTTTGGATCAAAGCAACGAAGAT------ 1431

Qy        493 ProIleIleThrSerSerLeuLeuGlnAspGlyIleThrThrThrGlnAlaGluIleSer 512
Db       1432 ----------------------------------------TCTGCTTCAAAGTCA 1446

Qy        513 ThrSerThrIleSerThrSerProThrAlaGlyAlaSerLeuHis------HisAlaHis 530
Db       1447 ACAAGTAACCAAGAAAAGAGCCCAATGGCGGACACTAATTACCATATGAATCCCAATCAC 1506

Qy        531 AsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrThrValAspHisIle 550
Db       1507 AACGGTGACCTAGAAGGCGTCACTGGAATGCAAGGAAGCCCCAAGAGACTAAGAACCAGC 1566

Qy        551 GluAsnAsnAlaLysLysGlnArgAsnAspMetHisLysPheSerProSerSerIleLeu 570
Db       1567 GACGAGACAATGATGCAGCCAATAAAT------GCGGATTTCAGCTCCAACGAGAAGCTC 1620

Qy        571 SerSerValAspMetGluAlaLysAlaArgGluSerSerAsnLysGlyPheThrAsnPro 590
Db       1621 ACGATGAAGATTCTAGAAGAACGGCAAGGGATAAGATCAGATGGTGGCTAC------CCT 1674

Qy        591 LeuMet---------AlaAlaTyrAlaMetGlyAspPheGlyArgPheAsp---ProHis 606
Db       1675 TTCATGGGTAATTTCGGGCAATACCAAATGGATGAGATGTCAAGATTTGATGTAGTCTCA 1734

Qy        607 AspGlnGlnMetThrAlaAsnPheHis------GlyAsnAsnGlyValSerLeuThrLeu 624
Db       1735 GACCAGGAGCTCATGGCGCAAAGGTACTCAGGAAACAACAATGGCGTGTCCCTCACGTTA 1794

Qy        625 GlyLeuProProSerGluAsnLeuAlaMetProValSerGln---------GlnAsnTyr 641
Db       1795 GGTTTACCTCATTGTGATAGCTTGTCGTCCACGCACCATCAGGGTTTCATGCAGACCCAC 1854

Qy        642 LeuSerAsnAspLeuGlySerArgSerGluMetGly----------------------- 653
Db       1855 CATGGGATTCCTATAGGGAGAAGAGTGAAAATAGGAGAAACAGAGGAATATGGACCCGCC 1914

Qy        654 ----------------------------------------------------SerHis 655
Db       1915 ACCATCAATGGTGGTAGCTCGACCACAACCGCACATTCATCAGCGGCAGCTGCCGCGGCT 1974

Qy        656 TyrAsnArgMetGlyTyrGluAsnIleAspPheGlnSerGlyAsnLysArgProThr 675
Db       1975 TACAATGGGATGAACATACAGAAC----------------CAGAAGAGATATGTGGCT 2016

Qy        676 GlnLeuLeuProAspPheVal 682
Db       2017 CAGTTATTGCCCGACTTCGTT 2037
```

RESULT 3
ADD30303
ID ADD30303 standard; cDNA; 2385 BP.
XX
AC ADD30303;
XX
DT 15-JAN-2004 (first entry)
XX
DE Plant yield-related polynucleotide clone G1589.
XX
KW ds; transcription factor; transgenic plant; growth rate; senescence;
KW seed germination rate; plant vigor; seedling vigor.
XX
OS Arabidopsis thaliana.
XX
PN WO2003013227-A2.
XX
PD 20-FEB-2003.
XX
PF 09-AUG-2002; 2002WO-US025805.
XX
PR 09-AUG-2001; 2001US-0310847P.
PR 19-NOV-2001; 2001US-0336049P.
PR 11-DEC-2001; 2001US-0338692P.
PR 14-JUN-2002; 2002US-00171468.
XX
PA (MEND-) MENDEL BIOTECHNOLOGY INC.
XX
PI Ratcliffe O, Riechmann JL, Adam LJ, Dubell AT, Heard JE;
PI Pilgrim ML, Jiang C, Reuber TL, Creelman RA, Pineda O, Yu G;
PI Broun PE;
XX
DR WPI; 2003-248221/24.
DR P-PSDB; ADD30304.
XX
PT New plant transcription factor polynucleotides and polypeptides, useful
PT in producing transgenic plants with commercially valuable properties,
PT such as an alteration in a plant growth characteristic, e.g. growth rate
PT or apomixis.
XX
PS Disclosure; SEQ ID NO 332; 454pp; English.
XX
CC The invention relates to a number of isolated Arabidopsis thaliana cDNA
CC sequences and their encoded proteins which are especially transcription
CC factor related cDNA's and proteins. The isolated or recombinant plant
CC transcription factor polynucleotides and polypeptides are useful in
CC producing transgenic plants with commercially valuable properties, i.e.
CC modified or altered desirable traits as compared to a reference plant,
CC such as an alteration in a plant growth characteristic, e.g. growth rate,

CC germination rate of seeds, vigor of plants and seedlings, or leaf and
CC flower senescence. Sequence information related to the polynucleotides
CC and polypeptides can also be used in bioinformatic search methods. The
CC transgenic plant is useful for growing a progeny plant from a parent
CC plant. This sequence represents one of the cDNAs of the invention.
XX
SQ Sequence 2385 BP; 748 A; 578 C; 565 G; 494 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.39e-85	Length:	2385
Score:	1148.50	Matches:	306
Percent Similarity:	51.67%	Conservative:	80
Best Local Similarity:	40.96%	Mismatches:	202
Query Match:	32.23%	Indels:	159
DB:	10	Gaps:	28

US-10-624-201A-2 (1-688) x ADD30303 (1-2385)

```
Qy         22 HisGlyAsn-----------------SerAsnAsnAsnAsnIleGlnThrLeuTyrLeu 35
               |||||||||                      ::::::     :::|||||||||   |||
Db        194 CACGGAAACCCACCGGAGATCTCTGCCGGATCCGACGGTGGTCTTCAAACGTTGATCCTC 253

Qy         36 MetAsnProAsnAsnTyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGlnGlnGln 55
               |||||||||      |||:::|||   ||||||      |||
Db        254 ATGAATCCAACTACTTACGTTCAG---TACACCCAACAAGACAACGACTCGAACAACAAC 310

Qy         56 LeuLeuPheLeuAsnSerSerProAlaAlaSerAsnAlaLeuCysHisAlaAsnIle--- 74
                        |||:::|||      ::::::|||       :::   |||
Db        311 -----------AACAACAGCAACAATAGCAACAACAACAACACAAACACAAACACAAAC 358

Qy         75 ------------------------------GlnHisAlaProLeu------GlnGlnGln 82
                                            |||||||||                 |||
Db        359 AACAACAACAGTAGTTTCGTTTTCCTCGATTCCCACGCGCCGCAGCCAAACGCGAGCCAG 418

Qy         83 HisPheValGlyValProLeu---------------ProAlaValSerLeuHisAspGln 97
                 |||||||||:::||||||                    :::::::::      |||
Db        419 CAGTTCGTCGGAATACCACTCTCAGGTCACGAAGCTGCTTCCATTACAGCCGCCGACAAC 478

Qy         98 IleAsn---HisHisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnSerGlnGln 116
              |||:::     ||||||          |||:::     :::          |||   |||
Db        479 ATCTCCGTACTTCACGGTTATCCTCCGCGCGTGCAGTACAGTCTCTACGGTAGCCACCAA 538

Qy        117 ValIleValProSerSerThrGlyValSerAlaThrSerCysGlyGlyIleThrThrAsp 136
              |||
Db        539 GTG------------------------------------------------------- 541

Qy        137 LeuAlaSerGlnLeuAlaPheGlnArgProIleProThrProGlnHisArgGln------ 154
                                                 |||   |||:::|||
Db        542 -----------------------------------GATCCCACTCACCAGCAAGCCGCG 565

Qy        155 ------GlnGlnGlnGlnGlnGlyGlyLeuSerLeuSerLeuSerProGlnLeuGlnGln 172
                    :::   |||   ||||||||||||:::||||||   |||   ||||||
Db        566 TGTGAGACGCCACGCGCGCAGCAAGGCCTCTCTTTAACCCTCTCGTCTCAACAGCAGCAG 625

Qy        173 GlnIleSerPheAsnAsn--------------AsnIleSerSerSerSerProArgThr 187
              |||         :::                 ::::::          |||
Db        626 CAACAGCAACATCATCAACAACACCAGCCTATTCACGTCGGATTCGGGTCCGGACATGGA 685

Qy        188 AsnAsnValThrIleArgGlyThrLeuAspGlySerSer--------------SerAsn 202
                 ::::::   :::                  ||||||                :::|||
Db        686 GAAGATATCCGGGTCGGGTCTGGCTCTACAGGATCGGGGGTAACAAACGGTATAGCTAAT 745

Qy        203 MetValLeuGlySerLysTyrLeuLysAlaAlaGlnGluLeuLeuAspGluValValAsn 222
              :::|||      ||||||||||||||||||||||||||||||||||||||||||||||||
Db        746 CTTGTT---AGCTCCAAGTACTTGAAGGCAGCACAAGAGCTTCTTGACGAAGTAGTCAAC 802

Qy        223 IleValGlyLysSerIleLysGlyAspAspGln-----------LysLysAspAsnSer 238
                          :::                ||||      ||||||   :::
Db        803 GCTGATTCCGATGACATGAACGCTAAATCCCAACTATTCTCATCGAAAAAGGGTAGTTGC 862

Qy        239 MetAsnLysGluSerMetProLeuAlaSer---AspValAsnThrAsnSerSerGlyGly 257
                 |||   :::  :::      :::|||                  |||||||||
Db        863 GGAAATGATAAACCTGTCGGAGAATCATCGGCCGGCGCTGGAGGAGAAGGTTCCGGTGGC 922

Qy        258 GlyGluSerSerSerArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGlnArg 277
               |||      ::::::   ::::::            |||||||||   ||||||:::|||
Db        923 GGAGCAGAAGCAGCCGGGAAACGT---------CCGGTGGAGCTAGGCACGGCAGAGAGA 973

Qy        278 GlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeuGluGluValGluGlnArg 297
              ||||||:::||||||||||||||||||||||      ||||||   ||||||||||||||||
Db        974 CAAGAAATACAGATGAAGAAAGCAAAACTTAGTAACATGCTTCATGAGGTGGAGCAGAGA 1033

Qy        298 TyrArgGlnTyrHisHisGlnMetGlnIleIleValLeuSerPheGluGlnValAlaGly 317
              ||||||||||||||   |||||||||::::::::   ||||||||||   ||||||
Db       1034 TATAGACAGTACCACCAGCAGATGCAGATGGTGATCTCTTCGTTCGAGCAAGCGGCAGGG 1093

Qy        318 IleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArg 337
              ||||||||||||||||||||||||   |||||||||      ||||||:::|||||||||
Db       1094 ATAGGATCAGCGAAGTCATACACGTCGCTAGCATTGAAAACCATATCAAGACAGTTCCGT 1153

Qy        338 CysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGluGlu 357
              ||||||||:::|||||||||   |||:::||||||   :::|||||||||||||||||||
Db       1154 TGCTTGAAAGAGGCGATCGCTGGTCAGATAAAAGCGGCCAACAAGAGTCTTGGGGAGGAA 1213

Qy        358 GluGlyLeuGly------GlyLysIleGluGlySerArgLeuLysPheValAspHisHis 375
              :::   :::          |||:::   ||||||||||||||||||||||||||||||||
Db       1214 GATTCAGTGTCTGGTGTTGGGAGGTTTGAGGGGTCGAGGCTCAAGTTCGTGGACCACCAC 1273

Qy        376 LeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGln---------ProAsnAla 392
              ||||||||||||||||||||||||||:::||||||:::|||             ||||||
Db       1274 TTGAGACAGCAAAGAGCTCTTCAACAACTGGGAATGATTCAACATCCTTCCAATAATGCT 1333

Qy        393 TrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeuPhe 412
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1334 TGGAGACCTCAACGTGGTCTCCCAGAACGAGCCGTCTCAGTTCTCCGTGCTTGGCTCTTC 1393

Qy        413 GluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGlnThr 432
              ||||||||||||||||||||||||||||||||||||||   ||||||||||||||||||
Db       1394 GAACACTTTCTTCATCCATACCCTAAGGATTCGGACAAGCACATGCTAGCTAAGCAAACA 1453

Qy        433 GlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLys 452
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1454 GGACTCACTCGTAGCCAGGTGTCGAACTGGTTTATAAACGCGAGAGTTCGGTTATGGAAA 1513

Qy        453 ProMetValGluGluMetTyrLeuGluGluValLysAsnGlnGluGlnAsnSerThrAsn 472
              ||||||||||||||||||:::||||||:::|||   |||   :::|||
Db       1514 CCAATGGTGGAGGAGATGTACATGGAGGAAATGAAGGAGCAGGCAAAGAACATG------ 1567

Qy        473 ThrSerGlyAspAsnLysAsnLysGluThrAsnIleSerAlaProAsnGluGluLysHis 492
                                  ::::::|||   :::          ||||||:::
Db       1568 ------------GGATCCATGGAAAAGACTCCTTTGGATCAAAGCAACGAAGAT------ 1609

Qy        493 ProIleIleThrSerSerLeuLeuGlnAspGlyIleThrThrThrGlnAlaGluIleSer 512
                                                          |||      |||
Db       1610 ------------------------------------------TCTGCTTCAAAGTCA 1624

Qy        513 ThrSerThrIleSerThrSerProThrAlaGlyAlaSerLeuHis------HisAlaHis 530
              ||||||          ||||||   |||      :::   |||          |||
Db       1625 ACAAGTAACCAAGAAAAGAGCCCAATGGCGGACACTAATTACCATATGAATCCCAATCAC 1684

Qy        531 AsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrThrValAspHisIle 550
              |||             |||          |||                    :::
Db       1685 AACGGTGACCTAGAAGGCGTCACTGGAATGCAAGGATGCCCCAAGAGACTAAGAACCAGC 1744

Qy        551 GluAsnAsnAlaLysLysGlnArgAsnAspMetHisLysPheSerProSerSerIleLeu 570
              :::          :::      |||           ||||||   :::      |||
Db       1745 GACGAGACAATGATGCAGCCAATAAAT------GCGGATTTCAGCTCCAACGAGAAGCTC 1798

Qy        571 SerSerValAspMetGluAlaLysAlaArgGluSerSerAsnLysGlyPheThrAsnPro 590
              :::       :::|||   :::           |||:::   |||:::      |||
Db       1799 ACGATGAAGATTCTAGAAGAACGGCAAGGGATAAGATCAGATGGTGGCTAC------CCT 1852

Qy        591 LeuMet---------AlaAlaTyrAlaMetGlyAspPheGlyArgPheAsp---ProHis 606
                 |||         |||  |||   :::      |||||||||
Db       1853 TTCATGGGTAATTTCGGGCAATACCAAATGGATGAGATGTCAAGATTTGATGTAGTCTCA 1912
```

```
Qy        607 AspGlnGlnMetThrAlaAsnPheHis------GlyAsnAsnGlyValSerLeuThrLeu 624
              ||||||::::::   |||      :::          ||||||||||||||||||||||
Db       1913 GACCAGGAGCTCATGGCGCAAAGGTACTCAGGAAACAACAATGGCGTGTCCCTCACGTTA 1972

Qy        625 GlyLeuProProSerGluAsnLeuAlaMetProValSerGln---------GlnAsnTyr 641
              |||||||||      ::::::|||:::          |||         |||   :::
Db       1973 GGTTTACCTCATTGTGATAGCTTGTCGTCCACGGACCATCAGGGTTTCATGCAGACCCAC 2032

Qy        642 LeuSerAsnAspLeuGlySerArgSerGluMetGly----------------------- 653
                      :::|||   |||   ::::::|||
Db       2033 CATGGGATTCCTATAGGGAGAAGAGTGAAAATAGGAGAAACAGAGGAATATGGACCCGCC 2092

Qy        654 ------------------------------------------------------SerHis 655
                                                                     :::
Db       2093 ACCATCAATGGTGGTAGCTCGACCACAACCGCACATTCATCAGCGGCAGCTGCCGCGGCT 2152

Qy        656 TyrAsnArgMetGlyTyrGluAsnIleAspPheGlnSerGlyAsnLysArgPheProThr 675
              ||||||   |||      :::|||                 ||||||:::
Db       2153 TACAATGGGATGAACATACAGAAC-----------------CAGAAGAGATATGTGGCT 2194

Qy        676 GlnLeuLeuProAspPheVal 682
              |||||||||||||||||||
Db       2195 CAGTTATTGCCCGACTTCGTT 2215
```

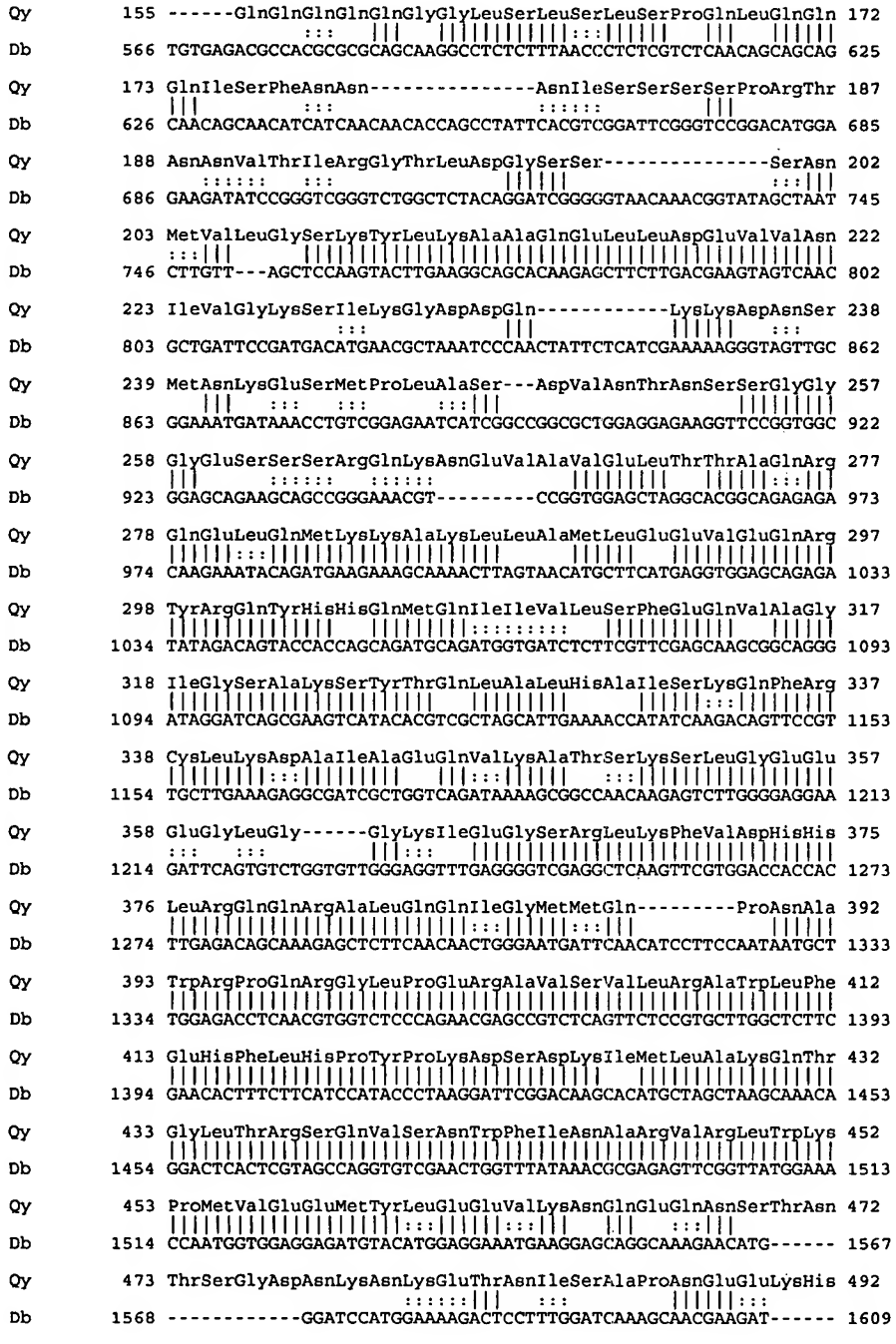
RESULT 4
ADI44208
ID ADI44208 standard; DNA; 2385 BP.
XX
AC ADI44208;
XX
DT 22-APR-2004 (first entry)
XX
DE Plant transcription factor related polynucleotide #1693.
XX
KW transgenic; plant; enhanced tolerance to abiotic stress;
KW glyphosphate tolerance; hormone sensitivity; disease resistance;
KW sugar sensing; flowering; flower structure; stem bifurcation;
KW branching pattern; apical dominance; trichome; stem morphology;
KW root growth; root hair; seed development; cell proliferation;
KW cell differentiation; premature senescence; necrosis; plant size;
KW leaf morphology; seed morphology; seed biochemistry; root anthocyanin;
KW plant anthocyanin; light response; shade avoidance; bioinformatic;
KW transcription factor; ds.
XX
OS Unidentified.
XX
PN US2004019927-A1.
XX
PD 29-JAN-2004.
XX
PF 25-FEB-2003; 2003US-00374780.
XX
PR 18-APR-2001; 2001US-00837944.
XX
PA (SHER/) SHERMAN B K.
PA (RIEC/) RIECHMANN J L.
PA (JIAN/) JIANG C.
PA (HEAR/) HEARD J E.
PA (HAAK/) HAAKE V.
PA (CREE/) CREELMAN R A.
PA (RATC/) RATCLIFFE O.
PA (ADAM/) ADAM L J.
PA (REUB/) REUBER T L.
PA (KEDD/) KEDDIE J.
PA (BROU/) BROUN P E.
PA (PILG/) PILGRIM M L.
PA (DUBE/) DUBELL A N.
PA (PINE/) PINEDA O.
PA (YUGG/) YU G.
XX
PI Sherman BK, Riechmann JL, Jiang C, Heard JE, Haake V;
PI Creelman RA, Ratcliffe O, Adam LJ, Reuber TL, Keddie J, Broun PE;
PI Pilgrim ML, Dubell AN, Pineda O, Yu G;
XX
DR WPI; 2004-132245/13.
DR P-PSDB; ADI44209.
XX
PT New transgenic plant comprising a recombinant polynucleotide of any one
PT of more than 500 nucleotide sequences, useful in bioinformatic search
PT methods.
XX
PS Disclosure; SEQ ID NO 2671; 435pp; English.
XX
CC The invention describes a transgenic plant comprising a recombinant
CC polynucleotide of any one of more than 500 nucleotide sequences fully
CC defined in the specification or its complement. The method of the
CC invention can be used to produced a plant having altered traits such as:
CC enhanced tolerance to abiotic stress; glyphosphate tolerance; hormone
CC sensitivity; disease resistance; sugar sensing; early or late flowering;
CC altered flower structure, change in stem bifurcations, altered branching
CC pattern, reduced apical dominance, reduced trichome density; lack of
CC trichomes; reduced ectopic trichome development; altered trichome
CC development; increase in trichome number; altered stem morphology;
CC increased root growth; increased root hairs; altered seed development;
CC altered cell proliferation or cell differentiation; rapid development;
CC premature senescence; increased necrosis; increase in seedling or plant
CC size; decreased plant size; leaf morphology; seed morphology; seed
CC biochemistry; increase in root anthocyanins; increase in plant
CC anthocyanins, or alteration in light response or shade avoidance. The
CC transgenic plant, polynucleotides and polypeptides are useful in
CC bioinformatic search methods. This sequence represents a plant
CC transcription factor related polynucleotide.
XX
SQ Sequence 2385 BP; 748 A; 578 C; 565 G; 494 T; 0 U; 0 Other;

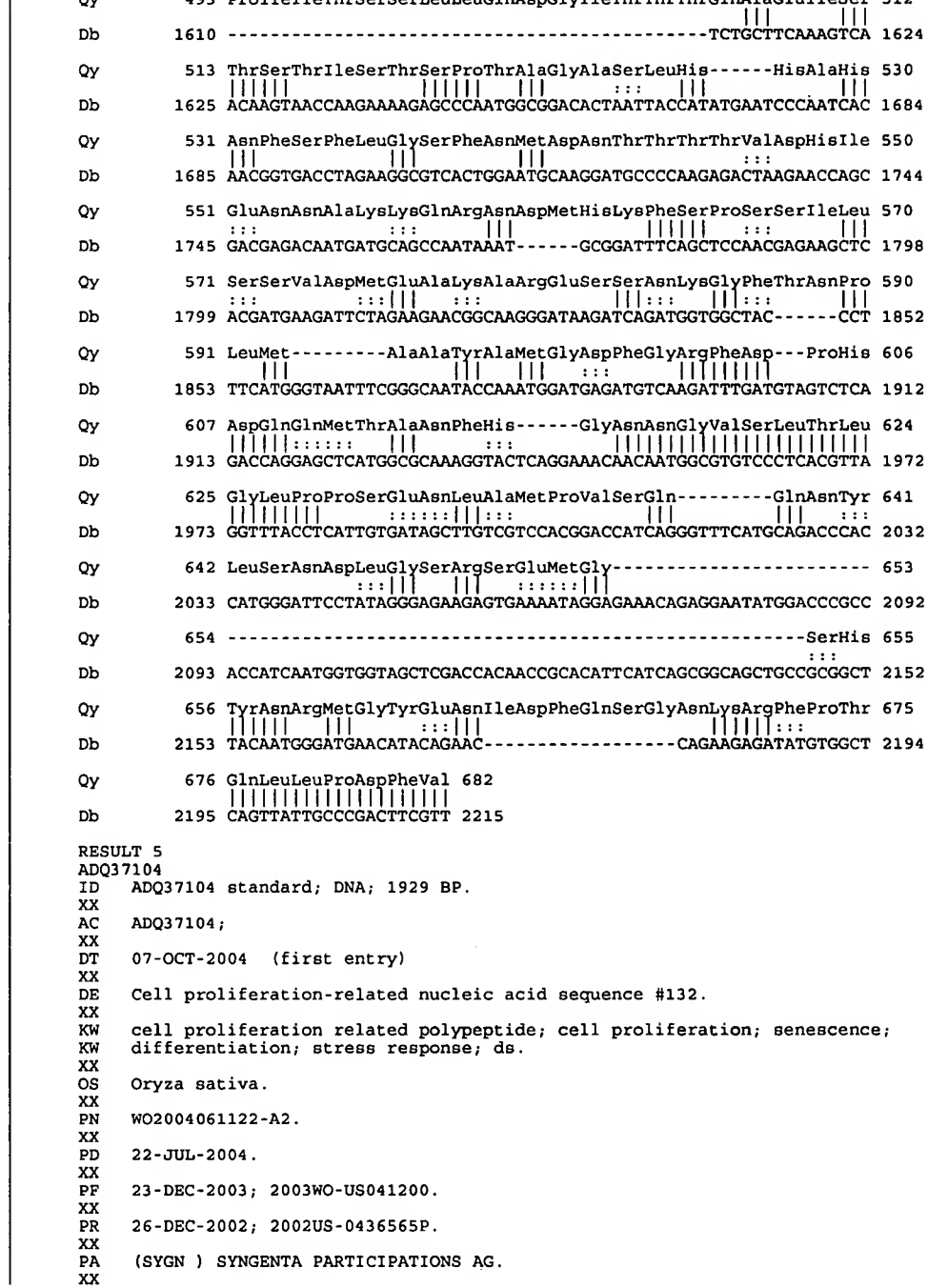
Alignment Scores:

Pred. No.:	2.39e-85	Length:	2385
Score:	1148.50	Matches:	306
Percent Similarity:	51.67%	Conservative:	80
Best Local Similarity:	40.96%	Mismatches:	202
Query Match:	32.23%	Indels:	159
DB:	12	Gaps:	28

US-10-624-201A-2 (1-688) x ADI44208 (1-2385)

```
Qy         22 HisGlyAsn-----------------SerAsnAsnAsnAsnIleGlnThrLeuTyrLeu 35
              |||||||||                 ::::::      :::|||||||||   |||
Db        194 CACGGAAACCCACCGGAGATCTCTGCCGGATCCGACGGTGGTCTTCAAACGTTGATCCTC 253

Qy         36 MetAsnProAsnAsnTyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGlnGlnGln 55
              |||||||||      |||:::|||   ||||||      |||
Db        254 ATGAATCCAACTACTTACGTTCAG---TACACCCAACAAGACAACGACTCGAACAACAAC 310

Qy         56 LeuLeuPheLeuAsnSerSerProAlaAlaSerAsnAlaLeuCysHisAlaAsnIle--- 74
                        |||:::|||      ::::::|||         :::   |||
Db        311 -----------AACAACAGCAACAATAGCAACAACAACAACACAAACACAAACACAAAC 358

Qy         75 ----------------------------GlnHisAlaProLeu------GlnGlnGln 82
                                           |||||||||               |||
Db        359 AACAACAACAGTAGTTTCGTTTTCCTCGATTCCCACGCGCCGCAGCCAAACGCGAGCCAG 418

Qy         83 HisPheValGlyValProLeu--------------ProAlaValSerLeuHisAspGln 97
              |||||||||:::||||||                 :::::::::      |||
Db        419 CAGTTCGTCGGAATACCACTCTCAGGTCACGAAGCTGCTTCCATTACAGCCGCCGACAAC 478

Qy         98 IleAsn---HisHisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnSerGlnGln 116
              |||:::      ||||||         |||:::      :::         |||   |||
Db        479 ATCTCCGTACTTCACGGTTATCCTCCGCGCGTGCAGTACAGTCTCTACGGTAGCCACCAA 538

Qy        117 ValIleValProSerSerThrGlyValSerAlaThrSerCysGlyGlyIleThrThrAsp 136
              |||
Db        539 GTG---------------------------------------------------------- 541

Qy        137 LeuAlaSerGlnLeuAlaPheGlnArgProIleProThrProGlnHisArgGln------ 154
                                                  |||   |||:::|||
Db        542 ----------------------------------GATCCCACTCACCAGCAAGCCGCG 565
```





```
RESULT 5
ADQ37104
ID   ADQ37104 standard; DNA; 1929 BP.
XX
AC   ADQ37104;
XX
DT   07-OCT-2004  (first entry)
XX
DE   Cell proliferation-related nucleic acid sequence #132.
XX
KW   cell proliferation related polypeptide; cell proliferation; senescence;
KW   differentiation; stress response; ds.
XX
OS   Oryza sativa.
XX
PN   WO2004061122-A2.
XX
PD   22-JUL-2004.
XX
PF   23-DEC-2003; 2003WO-US041200.
XX
PR   26-DEC-2002; 2002US-0436565P.
XX
PA   (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
```

PI Cooper B;
XX
DR WPI; 2004-534388/51.
XX
PT New nucleic acid molecule encoding a cell proliferation-related
PT polypeptide, useful for modulating cell proliferation, senescence,
PT differentiation, development, and stress response in plants, and for
PT producing enhanced food crops.
XX
PS Claim 57; SEQ ID NO 263; 408pp; English.
XX
CC The present invention relates to an isolated nucleic acid molecule
CC encoding a cell proliferation-related polypeptide. The nucleic acid
CC molecule and the encoded polypeptide, and methods are useful for
CC modulating cell proliferation, senescence, differentiation, development,
CC and stress response in plants, and for producing enhanced food crops. The
CC present sequence represents a cell proliferation-related nucleic acid
CC sequence. The present sequence is published separately from the main body
CC of the specification as EPO data.
XX
SQ Sequence 1929 BP; 575 A; 459 C; 463 G; 432 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	5.03e-67	Length:	1929
Score:	926.50	Matches:	262
Percent Similarity:	51.52%	Conservative:	110
Best Local Similarity:	36.29%	Mismatches:	225
Query Match:	26.00%	Indels:	125
DB:	12	Gaps:	31

US-10-624-201A-2 (1-688) x ADQ37104 (1-1929)

```
Qy          2 TyrTyrGlnGlyThrSerAspAsnThrAsnIleGlnAlaAspHisGlnGlnArgHisAsn 21
              ||||||          :::      :::   ||||||   :::          :::
Db         10 TACTACTCGAGCCCTGGCAATGAAAGGGACTCGCAAGCTATGTACCCAGCG------GAT 63

Qy         22 HisGlyAsnSerAsnAsnAsnAsnIleGlnThrLeu---TyrLeuMetAsnProAsnAsn 40
                |||||||||:::                  :::      ::::::   |||   |||
Db         64 TCAGGCAATTCATCATATCCTGTGCCATCAGCAATAGGAAACATGTTATATCCTGGCAAT 123

Qy         41 TyrMetGlnGly---TyrThr------ThrSerAspThrGlnGlnGlnGlnGlnLeuLeu 57
                     |||   ||||||      :::          |||   ||||||      :::
Db        124 GGGTCTTCTGGGCCATACACGGAATTCAGTGGCATTATCCAGCATCAGCAGAATTTCATG 183

Qy         58 PheLeuAsnSerSerProAlaAla---------SerAsnAlaLeuCysHisAlaAsnIle 74
                 |||         |||   |||             ::::::              |||:::
Db        184 GAGCTGCCTGGCCATCCAACTGCGATCTCTCAAGATTCATCGTCACGGGAACCTAACATG 243

Qy         75 GlnHisAlaProLeuGlnGlnGlnHisPheValGlyValProLeuProAlaValSerLeu 94
                 :::   :::   |||:::   |||                ||||||      :::
Db        244 GTCGCGTCGTACATGGATCAGCGCTCTTTTGGA-----------CCTGCCAAAGATATG 291

Qy         95 HisAspGlnIleAsnHisHisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnSer 114
                 ::::::                ::::||      :::   :::          :::
Db        292 AGAAATGAG--------------ATGTTGATGCATCTGATGGATGGAGCACATAATGCT 336

Qy        115 GlnGlnValIleValProSerSerThrGlyValSerAlaThrSerCysGlyGlyIleThr 134
                   ::::::   :::   |||      ||||||           |||:::
Db        337 GGTGCTGATCTCATCCACAATGACACTCATAGCAGCGCGCAGATTGAGTTTGGCCTATTG 396

Qy        135 ThrAspLeuAlaSerGlnLeuAlaPheGlnArgProIleProThrProGlnHisArgGln 154
                 :::      |||                  |||   |||
Db        397 AACAACCACAATTCGATGAGCGTT-----------GCACCAGCACCAGGCCAA------ 438

Qy        155 GlnGlnGlnGlnGlnGlyGlyLeuSerLeuSerLeuSerProGlnLeu----------- 170
                                |||||||||||||||||::      :::
Db        439 -----------------GGATTGTCTCTGAGCCTCAACACGCATATCCTGGCGCCTTCG 480

Qy        171 -----------------GlnGlnGlnIleSerPheAsnAsnAsnIleSerSerSerSer 184
                              :::  ::::::      ::::::
Db        481 TATCCATACTGGTCTGCGAAAACAGAGTTGCTAACACCACACTCTTACCATGGTGATGAC 540

Qy        185 ProArgThrAsnAsnValThrIleArgGlyThrLeuAspGlySerSerSerAsnMetVal 204
                   |||      |||                  :::          :::|||      :::
Db        541 AACAGAATGAAGAAT-----------------ATGCAATCTGAGGCCTCACAGGCAATC 582

Qy        205 LeuGlySerLysTyrLeuLysAlaAlaGlnGluLeuLeuAspGluValValAsnIleVal 224
                  ||||||||||||||||||||||||||||||||||||||||||||||:::   |||
Db        583 AGAAACTCAAAGTATCTGAAAGCAGCACAAGAATTGCTTGATGAGGTCGTCAGT---GTT 639

Qy        225 GlyLysSerIleLysGlyAspAspGlnLysLysAsp------------AsnSerMetAsn 240
                  ||||||||||||      ||||||                        |||   |||
Db        640 TGGAAGAGTATAAAGCAGAAGGCTCAGAAAGACCAGGCTGAAGCAGGAAAATCAGATAAC 699

Qy        241 LysGluSerMetProLeuAlaSerAspValAsnThrAsnSerSerGlyGlyGlyGluSer 260
                  ||||||:::                              |||   |||   |||   |||
Db        700 AAAGAAGCC--------------------GAGGGGGGTTCGAAAGGTGAGGGGGTATCT 738

Qy        261 SerSerArgGlnLysAsn------GluValAlaValGluLeuThrThrAlaGlnArgGln 278
                  |||:::   |||::::::            |||   |||::::::   |||:::::|||
Db        739 TCCAACCCACAGGAGTCTACTGCCAATGCTGCACCAGAGATTTCTGCTGCTGAGAAACAA 798

Qy        279 GluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeuGluGluValGluGlnArgTyr 298
                  ||||||||   |||   ||||||||||:::|||||||||:::||||||:::::::::|||
Db        799 GAGCTCCAGAATAAGATGGCAAAACTTATGGCCATGTTGGATGAGGTGGACCGAAAGTAC 858

Qy        299 ArgGlnTyrHisHisGlnMetGlnIleIleValLeuSerPheGluGlnValAlaGlyIle 318
                  :::   |||:::||||||||||||||:::|||   ||||||:::   ||||||||
Db        859 AAACACTATTATCATCAAATGCAAATTGTAGTCTCATCTTTTGATATGGTTGCTGGGTCT 918

Qy        319 GlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArgCys 338
                  |||:::||||||   ||||||   :::|||||||      |||||||||   |||||||||
Db        919 GGAGCTGCCAAGCCTTATACTGCAGTGGCCCTTCAGACAATCTCAAAACATTTCAGATGT 978

Qy        339 LeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGluGluGlu 358
                  ||||||||||||||   :::|||:::              |||   ||||||||||||||
Db        979 CTGAAAGATGCTATCAACGATCAGATCAATGTTATCCGGAAGAAACTTGGAGAGGAGGAG 1038

Qy        359 GlyLeuGlyGlyLysIleGluGly------SerArgLeuLysPheValAspHisHisLeu 376
                  ||||||   ||||||      :::||||||:::::::::|||      |||
Db       1039 AGTTCATCTGGCAAA---GAGGGAAAATTAACGCGCCTCCGTTATATTGACCAGCAATTA 1095

Qy        377 ArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGlnProAsnAlaTrpArgProGln 396
                  ||||||||||||||   ||||||   |||::::::|||   ||||||||||||||||||
Db       1096 AGACAACAGCGCGCTTTCCAGCAGTATGGTTTGTTACAGCAAAATGCTTGGAGGCCACAG 1155

Qy        397 ArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeuPheGluHisPheLeu 416
                  ||||||||||||||   :::||||||:::||||||||||||||||||||||||||||||
Db       1156 AGGGGACTGCCCGAAAACTCAGTTTCAATTCTTCGTGCTTGGCTTTTTGAACACTTCCTT 1215

Qy        417 HisProTyrProLysAspSerAspLysIleMetLeuAlaLysGlnThrGlyLeuThrArg 436
                  ||||||||||||||||||||:::|||:::|||||||||:::||||||||||||||||||
Db       1216 CACCCGTATCCAAAAGATTCAGAAAAGCTGATGCTAGCGAGACAAACTGGCTTAACAAGA 1275

Qy        437 SerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLysProMetValGlu 456
                  ||||||:::||||||||||||||||||||||||||||||||||||||||||:::|||
Db       1276 AGTCAGATTTCAAATTGGTTCATAAATGCCCGTGTCCGCCTGTGGAAACCGATGATCGAA 1335

Qy        457 GluMetTyrLeuGluGluValLysAsnGlnGluGlnAsnSerThrAsnThrSerGlyAsp 476
                  :::||||||   |||||:::          :::   :::|||   |||:::|||   |||
Db       1336 GACATGTATAAAGAAGAGATTGGGGAGGCGGATCTCGACTCG---AACTCCTCCTCCGAC 1392

Qy        477 AsnLysAsnLysGluThrAsnIleSerAlaProAsnGluGluLysHisProIleIleThr 496
                  |||      :::      :::      |||   :::|||:::|||          :::
Db       1393 AACGTACCAAGGAGCAAGGACAAAATAGCAACATCTGAAGATAAGGAAGAT---CTGAAA 1449

Qy        497 SerSerLeuLeuGln-----------AspGlyIleThrThrThrGlnAlaGluIleSer 512
                  ||||||:::   |||                ::::      ::::::|||   |||
Db       1450 AGCTCTATGAGCCAGACTTATCAACCAAGCCAATTAGGTGAATCCAAAGCCAACATCGGG 1509

Qy        513 ThrSerThrIleSerThrSerProThrAlaGlyAlaSerLeuHisHisAlaHisAsn--- 531
                  ::::::      :::|||          |||       |||:::       |||
Db       1510 ATGATGAGCCTTGGTGGGGCACCA---------GCCGGCTTCCACAACGAAGGCAACCAG 1560

Qy        532 ---PheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrThrValAspHisIle 550
                     ||||||:::
```

```
Db       1561 GATGACAGCTTCATGAACCTAATG---------------------------------- 1584

Qy        551 GluAsnAsnAlaLysLysGlnArgAsnAspMetHisLysPheSerProSerSerIleLeu 570
                  |||  ||||||                    |||
Db       1585 ---------CTGAAGGACCAAAGG--------------------CCGGGCGAAGCCGAA 1614

Qy        571 SerSerValAspMetGluAlaLysAlaArgGluSerSerAsnLysGlyPheThrAsnPro 590
               |||:::      :::|||   |||      |||                  |||
Db       1615 GGCAGCCTCCTCCATGATGCCGTCGCCCATCATTCCGATGAG-----------AACGCT 1662

Qy        591 LeuMetAlaAlaTyrAlaMetGlyAspPheGlyArgPheAspProHisAspGlnGlnMet 610
                  ||||||   :::         ||||||:::
Db       1663 CGGTTCATGGCTTACCATTTGTCGGGGCTCGGAAGATAC--------------------- 1701

Qy        611 ThrAlaAsnPheHisGlyAsnAsnGlyValSerLeuThrLeuGlyLeuPro---ProSer 629
                  ||||||:::  ||||||||||||||||||||      |||
Db       1702 --------------GGGAACAGCAATGTGTCATTGACACTTGGCTTACAGCATCCTGAC 1746

Qy        630 GluAsnLeuAlaMetProValSerGlnGlnAsnTyrLeuSerAsnAspLeuGlySerArg 649
               |||::::::        |||      :::                ||:::
Db       1747 AACAGGCTTTCGGTA---------CAGAACACTCATCAGCCAGGTTTCGCCGGTGCTGGA 1797

Qy        650 SerGluMetGlySerHisTyrAsnArgMetGlyTyr---------GluAsnIleAspPhe 666
               |||:::   :::            :::|||                :::   |||:::
Db       1798 GAAGAAATTTACAACTCCACCGCTTCTCTCGGTGTCGCTGCAGCTTCCTCTTCGGACTAC 1857

Qy        667 GlnSerGlyAsn--------------LysArgPhe---ProThrGlnLeuLeuProAsp 680
               :::|||   |||            :::||||||   |||:::   |||:::   |||
Db       1858 GAATCCACGAACCAAATAGATCAAAGGCAACGGTTCGAACCATCGCCTCTAATGCATGAT 1917

Qy        681 PheVal 682
               ||||||
Db       1918 TTTGTG 1923
```

RESULT 6
ADO62706
ID ADO62706 standard; DNA; 1971 BP.
XX
AC ADO62706;
XX
DT 15-JUL-2004 (first entry)
XX
DE Transcription factor G2550 orthologous sequence, SEQ ID 1173.
XX
KW Plant; transcription factor; transgenic plant; abiotic stress tolerance;
KW osmotic stress tolerance; cold tolerance; heat tolerance;
KW low nitrogen tolerance; low phosphate tolerance; fungal disease;
KW glyphosate resistance; flowering; fertility; seed development; ds.
XX
OS Glycine max.
XX
PN WO2004031349-A2.
XX
PD 15-APR-2004.
XX
PF 18-SEP-2003; 2003WO-US030292.
XX
PR 18-SEP-2002; 2002US-0411837P.
PR 17-DEC-2002; 2002US-0434166P.
PR 24-APR-2003; 2003US-0465809P.
XX
PA (MEND-) MENDEL BIOTECHNOLOGY INC.
XX
PI Jiang C, Heard JE, Ratcliffe O, Creelman RA, Adam LJ, Reuber TL;
PI Riechmann JL, Haake V, Dubell AN, Keddie JS, Sherman BK;
XX
DR WPI; 2004-330163/30.
XX
PT New recombinant polynucleotide encoding transcription factor
PT polypeptides, useful for producing transgenic plants with advantageous
PT properties compared to a reference plant.
XX
PS Claim 1; SEQ ID NO 1173; 510pp; English.
XX
CC The present invention relates to novel plant transcription factor
CC proteins (I) and nucleotide sequences (II) (ADO61534-ADO63778). The
CC sequences can be used to produce transgenic plants, which overexpress
CC (II), where the transgenic plant has an altered trait as compared to a
CC non-transgenic plant or wild-type plant. The transgenic plant comprises
CC an altered trait selected from increased tolerance to abiotic stress,
CC increased tolerance to osmotic stress, increased tolerance to cold,
CC increased germination in cold, increased tolerance to heat, increased
CC germination in heat, increased tolerance to freezing conditions,
CC increased tolerance to low nitrogen conditions, increased tolerance to
CC low phosphate conditions, increased tolerance to disease, including
CC fungal disease and particularly Erysiphe, Fusarium and Botrytis,
CC increased tolerance to multiple fungal pathogens, increased resistance to
CC glyphosate, increased sensitivity to ABA, reduced sensitivity to ABA,
CC increased sensitivity to ACC, altered sugar sensing, increased tolerance
CC to sugars, altered carbon/nitrogen sensing, early flowering, late
CC flowering, altered flower structure, loss of flower determinacy, reduced
CC fertility, altered shoot meristem development, altered branching pattern,
CC altered stem morphology, altered vascular tissue structure, reduced
CC apical dominance, altered trichome density, altered trichome development,
CC altered trichome structure, altered root development, altered shade
CC avoidance, altered seed development, altered seed ripening, altered seed
CC germination, slow growth, fast growth, altered cell differentiation,
CC altered cell proliferation, altered cell expansion, altered phase change,
CC altered senescence, abnormal embryo development, altered programmed cell
CC death, lethality when overexpressed, altered necrosis patterns, increased
CC plant size, increased biomass, large seedlings, dwarfed plants, dark
CC green leaves, change in leaf shape, increased leaf size and mass, light
CC green or gray leaves, glossy leaves, altered abaxial/adaxial polarity,
CC altered seed coloration, altered seed size, altered seed shape, large
CC seed, increased leaf wax, increased leaf fatty acids, altered seed oil
CC content, altered seed protein content, altered seedprenyl content,
CC altered leaf prenyl lipid content, increased anthocyanin levels, and
CC decreased anthocyanin levels. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1971 BP; 667 A; 360 C; 451 G; 493 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.16e-58	Length:	1971
Score:	826.00	Matches:	216
Percent Similarity:	52.88%	Conservative:	78
Best Local Similarity:	38.85%	Mismatches:	148
Query Match:	23.18%	Indels:	114
DB:	12	Gaps:	18

US-10-624-201A-2 (1-688) x ADO62706 (1-1971)

```
Qy        119 ValProSerSerThrGlyValSerAlaThrSerCysGlyGlyIleThrThrAspLeuAla 138
                 ||||||:::      |||   |||:::                :::      |||   |||
Db         19 GTCCCAAATTTAAAGGGAACTTCATCT--------------CTTAAAGATGATGAGGCC 63

Qy        139 SerGlnLeuAlaPheGlnArgProIleProThrProGlnHisArgGlnGlnGlnGlnGln 158
                 |||           ||||||   :::         :::
Db         64 AGC---------TTACAAAGGGAATTGAGAAATGCTGAGTGCATGGCATCTTTAGCTTCT 114

Qy        159 GlnGlyGlyLeuSerLeuSerLeuSerProGlnLeuGlnGlnGlnIleSerPheAsnAsn 178
                    ||||||                                          |||:::
Db        115 TCTGGTGGC-----------------------------------------TTTCACAAA 132

Qy        179 AsnIleSerSerSerSerProArgThrAsnAsnValThrIle---ArgGlyThrLeuAsp 197
                          :::|||:::      ::::::   :::        |||
Db        133 AGGGATGGTTTGTATAATCCACAACACCCCTCAATGTGCCTTGGTGAAGGCCAAAGTCAT 192

Qy        198 GlySer------SerSerAsnMetValLeuGlySerLysTyrLeuLysAlaAlaGlnGlu 215
                 ||||||      |||||||||   ||||||   |||:::|||||||||||||||||||
Db        193 GGATCACAAGGCTCGAGCAACAATGTCTTGAACTCACAATACCTCAAGGCAGCACAGGAG 252

Qy        216 LeuLeuAspGluValValAsnIleValGlyLysSerIleLysGlyAspAspGlnLysLys 235
                 ||||||||||||:::||||||   |||   |||::::::|||            :::|||
```

```
Db        253 TTGCTTGATGAAATAGTAAAT---GTCCGAAAGGCTTTGAAGCAAACTGGTTTGGAAAAG 309

Qy        236 AspAsnSerMetAsnLysGluSerMetProLeuAlaSerAspValAsnThrAsnSerSer 255
              |||                 :::            |||   :::   :::|||
Db        310 CAACAGAGTTTCCGTGACACTGGTTTA-----------GATGGCTCCAAAGATTCTGAT 357

Qy        256 GlyGlyGlyGluSerSerSerArgGlnLysAsn-----------------------Glu 267
              |||          |||   |||   |||   :::
Db        358 GGAAAATCTACCAGCCAATCTGTGCAGATGTCTTCAGGCCCCAATGGTTCCGCTGCTAAC 417

Qy        268 ValAlaValGluLeuThrThrAlaGlnArgGlnGluLeuGlnMetLysLysAlaLysLeu 287
                 :::   ||||||:::   |||:::||||||   |||      ||||||   ||||||
Db        418 GCTTCTTGTGAGCTATCACCTGCAGAACGGCAGAACTTGTTGGACAAGAAGACAAAGCTT 477

Qy        288 LeuAlaMetLeuGluGluValGluGlnArgTyrArgGlnTyrHisHisGlnMetGlnIle 307
              |||:::|||||||:::||||||:::::|||||||||||||||   ||||||||||||||
Db        478 TTGTCCATGCTGGATGAGGTGGATAAAAGATACAGACAGTACTGCCATCAGATGCAGATT 537

Qy        308 IleValLeuSerPheGluGlnValAlaGlyIleGlySerAlaLysSerTyrThrGlnLeu 327
              :::|||   ||||||:::   |||||||||   |||:::|||:::   ||||||   |||
Db        538 GTGGTGTCATCTTTTGACATGGTTGCTGGCTGTGGAGCAGCAGAACCATATACCACACTT 597

Qy        328 AlaLeuHisAlaIleSerLysGlnPheArgCysLeuLysAspAlaIleAlaGluGlnVal 347
              ||||||      ||||||:::   |||||||||||||:::|||||||||:::   |||:::
Db        598 GCCTTAAGAACAATTTCTCGCCACTTTCGGTGTTTGCGTGATGCCATCAGTGGCCAAATT 657

Qy        348 LysAlaThrSerLysSerLeuGlyGluGluGluGlyLeuGlyGlyLysIleGluGlySer 367
              :::   |||   :::||||||||||||:::||||||:::
Db        658 CAGGTGACCCAAAGAAGCCTTGGGGAGCAAGAGGGAATA-----------------CCC 699

Qy        368 ArgLeuLysPheValAspHisHisLeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMet 387
              ||||||:::::||||||      ||||||||||||:::||||||||||||:::|||:::
Db        700 CGTCTCCGCTATGTGGATCAGCAACTTAGACAACAAAAGGCCCTTCAGCAACTTGGTGTA 759

Qy        388 MetGlnProAsnAlaTrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeu 407
              |||:::      ||||||||||||||||||||||||||||||   :::||||||||||||
Db        760 ATGAGA---CAAGCTTGGAGGCCTCAGAGGGGACTTCCTGAAACCTCTGTTTCAGTACTC 816

Qy        408 ArgAlaTrpLeuPheGluHisPheLeuHisProTyrProLysAspSerAspLysIleMet 427
              ||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||
Db        817 CGTGCTTGGCTCTTTGAGCATTTCCTTCATCCTTATCCTAAGGATTCAGAGAAAATTATG 876

Qy        428 LeuAlaLysGlnThrGlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArg 447
              ||||||:::||||||||||||||||||:::||||||:::||||||||||||||||||||
Db        877 CTAGCAAGGCAAACTGGCTTAACAAGAAACCAGGTGGCAAACTGGTTCATTAATGCAAGG 936

Qy        448 ValArgLeuTrpLysProMetValGluGluMetTyrLeuGluGluValLysAsnGlnGlu 467
              ||||||||||||||||||||||||||||||||   ||||||      :::   |||
Db        937 GTGCGTCTATGGAAGCCAATGGTTGAGGAAATGTACAAAGAAGAATTTGGTGATTCTGAG 996

Qy        468 GlnAsnSerThrAsnThrSer------------------------------------ 474
              :::|||        |||
Db        997 ATGAGCAGCAATCTATTATCATCAGAGAAAACACTCAAAGCTCCAAGAGATGATGTTCAA 1056

Qy        475 ---GlyAspAsnLysAsnLysGluThrAsn---------IleSerAlaProAsnGluGlu 490
                 |||||||||   :::|||:::            |||:::      :::
Db       1057 GCTTCTGACAATAAAAGGGAAGAGTCCCAGGACAACTTAATAAACGTTGATGATAGTGTT 1116

Qy        491 LysHis---------ProIleIleThrSerSerLeuLeuGlnAspGlyIleThrThrThr 507
              :::|||            :::      :::|||   |||      ||||||   ::::::
Db       1117 CAGCATCATGGGTTGAAGTTAGATCATGCCTCGGAATTGGACAGAGGGATCCAAAGCAGT 1176

Qy        508 ------------------------------------------------GlnAlaGlu 510
                                                             |||
Db       1177 GATCATGGAGAAAATGCCATGGATCCTAGAATTGGAAAATTGCAAGGTGACCAAAGGTTC 1236

Qy        511 IleSerThrSerThrIleSerThrSerProThrAlaGly--------------AlaSer 525
                   ::::::   :::   ||||||      |||                ||||||
Db       1237 AACATGAACAACAGCAATAACAATAGTCCTTATTATGGTGATGGCTGCATAATGGCTTCT 1296

Qy        526 LeuHisHisAlaHisAsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThr 545
                    ::::::         |||||:::      :::|||         :::
Db       1297 ACTCCTGCCACATATGATTTACCAGAGTTAGGTAACATTGCCGTCGATGGCCACGTGTCG 1356

Qy        546 ThrValAspHisIleGluAsnAsnAlaLysLys------------GlnArgAsnAspMet 561
                       :::   |||         :::                  :::||||||
Db       1357 CTTGCATTGGAATTGAGGAACTGTGAAAGTCAAGGATTTGGTGTGTCAAATGATGACATG 1416

Qy        562 HisLysPheSerProSerSerIleLeuSerSerValAspMetGluAlaLysAlaArgGlu 581
              ||||||           ::::::   ||||||   :::   :::
Db       1417 CATAAACGACATAAGAAAACATTGGCTTCTTCCCCAGAGACTGAT--------------- 1461

Qy        582 SerSerAsnLysGlyPheThrAsnProLeuMetAlaAlaTyrAlaMetGlyAspPheGly 601
                                                          :::   |||:::
Db       1462 ----------------------------------------------TTGCTAGATTACCAT 1476

Qy        602 ArgPheAspProHisAspGlnGlnMetThrAlaAsnPheHisGlyAsn 617
                  ||||||      ||||||         |||      ||||||
Db       1477 TTCACAGACCCAGGAAAGCAACAA---------AACAAGTTTGGCAAT 1515

RESULT 7
ADO63042
ID   ADO63042 standard; DNA; 1617 BP.
XX
AC   ADO63042;
XX
DT   15-JUL-2004  (first entry)
XX
DE   Transcription factor G2546 coding sequence, SEQ ID 1509.
XX
KW   Plant; transcription factor; transgenic plant; abiotic stress tolerance;
KW   osmotic stress tolerance; cold tolerance; heat tolerance;
KW   low nitrogen tolerance; low phosphate tolerance; fungal disease;
KW   glyphosate resistance; flowering; fertility; seed development; ds.
XX
OS   Arabidopsis thaliana.
XX
PN   WO2004031349-A2.
XX
PD   15-APR-2004.
XX
PF   18-SEP-2003; 2003WO-US030292.
XX
PR   18-SEP-2002; 2002US-0411837P.
PR   17-DEC-2002; 2002US-0434166P.
PR   24-APR-2003; 2003US-0465809P.
XX
PA   (MEND-) MENDEL BIOTECHNOLOGY INC.
XX
PI   Jiang C,  Heard JE,  Ratcliffe O,  Creelman RA,  Adam LJ,  Reuber TL;
PI   Riechmann JL,  Haake V,  Dubell AN,  Keddie JS,  Sherman BK;
XX
DR   WPI; 2004-330163/30.
DR   P-PSDB; ADO63043.
XX
PT   New recombinant polynucleotide encoding transcription factor
PT   polypeptides, useful for producing transgenic plants with advantageous
PT   properties compared to a reference plant.
XX
PS   Claim 1; SEQ ID NO 1509; 510pp; English.
XX
CC   The present invention relates to novel plant transcription factor
CC   proteins (I) and nucleotide sequences (II) (ADO61534-ADO63778). The
CC   sequences can be used to produce transgenic plants, which overexpress
CC   (II), where the transgenic plant has an altered trait as compared to a
CC   non-transgenic plant or wild-type plant. The transgenic plant comprises
CC   an altered trait selected from increased tolerance to abiotic stress,
CC   increased tolerance to osmotic stress, increased tolerance to cold,
CC   increased germination in cold, increased tolerance to heat, increased
CC   germination in heat, increased tolerance to freezing conditions,
CC   increased tolerance to low nitrogen conditions, increased tolerance to
CC   low phosphate conditions, increased tolerance to disease, including
CC   fungal disease and particularly Erysiphe, Fusarium and Botrytis,
CC   increased tolerance to multiple fungal pathogens, increased resistance to
CC   glyphosate, increased sensitivity to ABA, reduced sensitivity to ABA,
CC   increased sensitivity to ACC, altered sugar sensing, increased tolerance
```

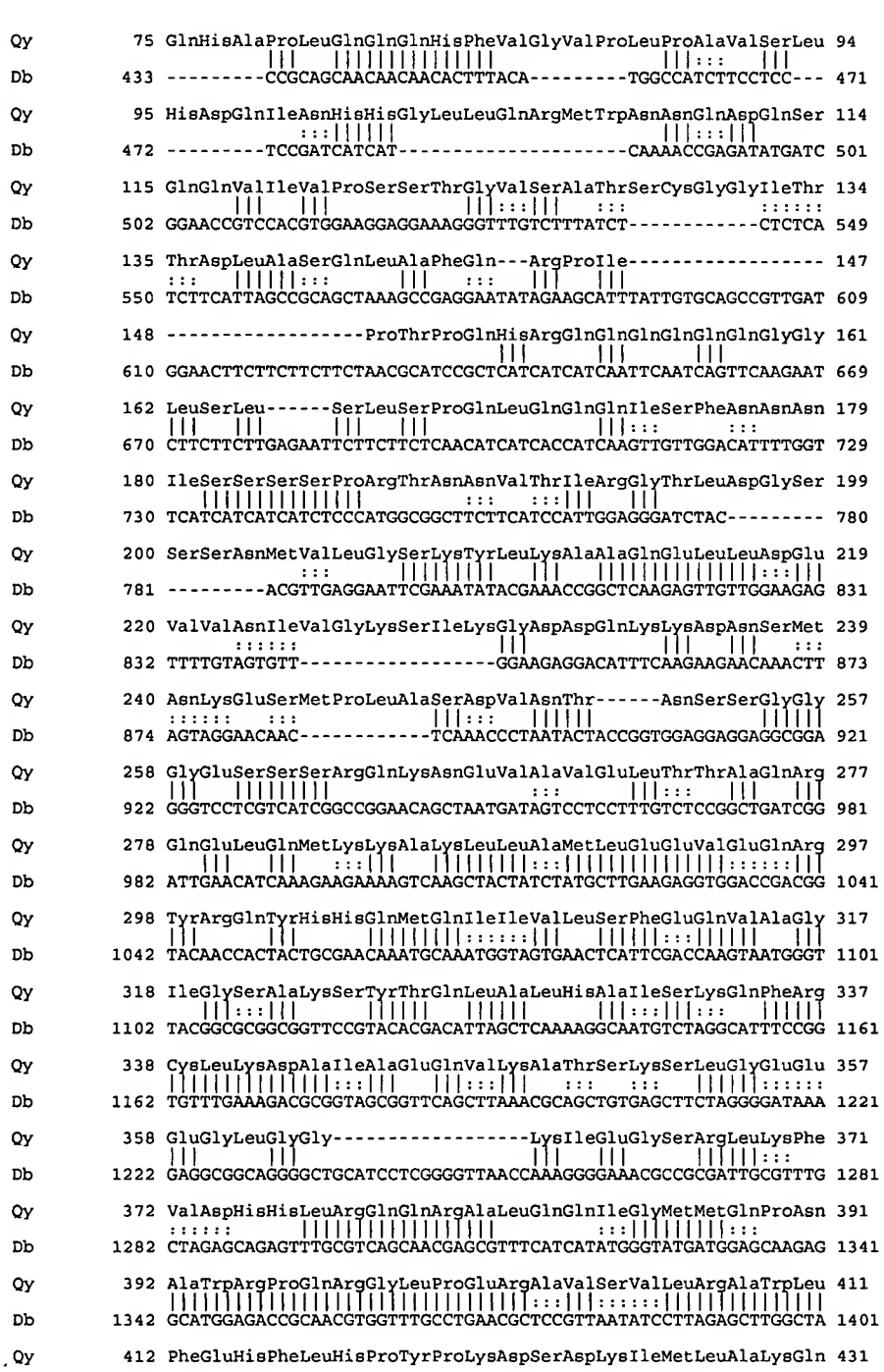
```
CC   to sugars, altered carbon/nitrogen sensing, early flowering, late
CC   flowering, altered flower structure, loss of flower determinacy, reduced
CC   fertility, altered shoot meristem development, altered branching pattern,
CC   altered stem morphology, altered vascular tissue structure, reduced
CC   apical dominance, altered trichome density, altered trichome development,
CC   altered trichome structure, altered root development, altered shade
CC   avoidance, altered seed development, altered seed ripening, altered seed
CC   germination, slow growth, fast growth, altered cell differentiation,
CC   altered cell proliferation, altered cell expansion, altered phase change,
CC   altered senescence, abnormal embryo development, altered programmed cell
CC   death, lethality when overexpressed, altered necrosis patterns, increased
CC   plant size, increased biomass, large seedlings, dwarfed plants, dark
CC   green leaves, change in leaf shape, increased leaf size and mass, light
CC   green or gray leaves, glossy leaves, altered abaxial/adaxial polarity,
CC   altered seed coloration, altered seed size, altered seed shape, large
CC   seed, increased leaf wax, increased leaf fatty acids, altered seed oil
CC   content, altered seed protein content, altered seedprenyl content,
CC   altered leaf prenyl lipid content, increased anthocyanin levels, and
CC   decreased anthocyanin levels. Note: The sequence data for this patent did
CC   not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1617 BP; 510 A; 335 C; 358 G; 414 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                1.18e-58      Length:          1617
Score:                    824.50        Matches:         212
Percent Similarity:       53.99%        Conservative:    79
Best Local Similarity:    39.33%        Mismatches:      111
Query Match:              23.14%        Indels:          137
DB:                       12            Gaps:            21

US-10-624-201A-2 (1-688) x ADO63042 (1-1617)

Qy        109 AsnAsnGlnAspGlnSerGlnGlnValIleValProSerSerThrGlyValSerAlaThr 128
              |||:::|||   |||:::||||||           ||||||         |||   |||
Db         58 AACCATCAACAACAAAACCAACAAGCT---TCTTCTTCATCCGCCGCCGCTTCTTTCACC 114

Qy        129 SerCysGlyGlyIleThrThrAspLeuAlaSerGlnLeuAlaPheGlnArgProIlePro 148
              ||||||   |||     :::   :::::::::   |||     |||
Db        115 GTCACCGGCGGCGATACT------GTTCGAAACGAGATGGTTTTTATCCCACCAACCACC 168

Qy        149 ThrProGln-----------------------------------------------HisArg 153
              |||                                                     |||
Db        169 ACAGGAGACGTTGTAACAGGAAACGGTACCGTTTCAAGCAGCGATCTAAGCTTTCACGAT 228

Qy        154 GlnGlnGlnGlnGlnGlyGlyLeuSerLeuSerLeuSerProGlnLeu--------- 170
                 |||              ||||||||||||||||||      |||:::
Db        229 GGTCAA--------------GGACTGTCTTTGAGCCTTGGTACTCAGATCTCTGTTGCT 273

Qy        171 ---------------GlnGlnGlnIleSerPheAsnAsn---------------AsnIle 180
                          |||   |||:::   |||   :::              ::::::
Db        274 CCGTTTCACTTTCATCAATACCAATTGGGGTTTACTAGTCAGAATCCTTCAATTTCAGTC 333

Qy        181 SerSerSerSerPro-----------------------ArgThrAsnAsnValThrIle 192
                  :::||||||                       ::::::      :::   :::
Db        334 AAGGAAACGTCACCGTTTCATGTGGATGAGATGAGTGTGAAGAGCAAAGAAATGATCTTG 393

Qy        193 ArgGlyThrLeuAspGlySerSerSer-------------------------------- 201
                 |||      |||   ||||||
Db        394 TTGGGTCAATCTGATCCTTCCTCTGGTTATGCTGGTAATGGTGGGAATGGCTTCTACAAC 453

Qy        202 ------------------------------------AsnMetValLeuGlySerLysTyr 209
                                                  :::   ||||||   |||:::|||
Db        454 AATTATCGGTATAATGAGACATCAGGAGGGTTTATGAGCAGCGTTCTGCGTTCTCGGTAT 513

Qy        210 LeuLysAlaAlaGlnGluLeuLeuAspGluValValAsnIleValGlyLysSerIleLys 229
              ||||||   ||||||   ||||||||||||||||||:::::
Db        514 CTTAAACCTGCTCAGAATTTGCTTGATGAAGTGGTTAGTGTC------------------ 555

Qy        230 GlyAspAspGlnLysLysAspAsnSerMetAsnLysGluSerMetProLeuAlaSerAsp 249
                      :::|||:::   |||   |||   |||:::   |||   :::   :::|||

Db        556 ---------AAGAAAGAACTAAACCAAATGGGGAAGAAGAAGATGAAAGTT---AATGAC 603

Qy        250 ValAsnThrAsnSer------------SerGlyGlyGlyGluSerSerSerArgGlnLys 265
                 |||:::   |||             ||||||||||||   ||||||
Db        604 TTTAACAGTGGTTCTAAGGAGATAGAAGGAGGAGGTGGTGAGTTATCGAGT---GATTCG 660

Qy        266 AsnGluValAlaValGluLeuThrThrAlaGlnArgGlnGluLeuGlnMetLysLysAla 285
              |||      ::::::||||||:::|||   :::|||:::|||||||||   ||||||
Db        661 AATGGGAAATCGATTGAGTTATCTACAATTGAACGTGAAGAGCTTCAGAACAAGAAGAAC 720

Qy        286 LysLeuLeuAlaMetLeuGluGluValGluGlnArgTyrArgGlnTyrHisHisGlnMet 305
              |||||||||   |||::::::||||||::::::||||||   ||||||:::|||||||||
Db        721 AAGCTTTTAACAATGGTTGATGAGGTAGATAAAAGATATAACCAATATTACCATCAAATG 780

Qy        306 GlnIleIleValLeuSerPheGluGlnValAlaGlyIleGlySerAlaLysSerTyrThr 325
              :::   :::       |||||||||   |||||||||:::||||||||||||   ||||||
Db        781 GAAGCATTAGCTTCATCATTTGAGATAGTAGCAGGACTTGGATCAGCTAAGCCTTACACA 840

Qy        326 GlnLeuAlaLeuHisAlaIleSerLysGlnPheArgCysLeuLysAspAlaIleAlaGlu 345
              :::||||||:::   ||||||:::   ||||||   |||:::|||||||||   |||
Db        841 TCAGTTGCTCTCAACAGAATCTCTCGCCATTTTCGTGCTCTTCGCGACGCAATAAAGGAA 900

Qy        346 GlnValLysAlaThrSerLysSerLeuGlyGlu-----------------GluGluGly 359
              |||::::::         :::   |||||||||                 :::|||
Db        901 CAAATTCAGATTGTTAGAGAAAAACTTGGGGAGAAAGGAGGAGAGTCGTTGGATGAGCAA 960

Qy        360 LeuGlyGlyLysIleGluGlySerArgLeuLysPheValAspHisHisLeuArgGlnGln 379
              |||   :::|||          ||||||:::::::::|||       ||||||||||
Db        961 CAAGGAGAGAGGATA------CCAAGGTTGAGGTATTTAGATCAACGGTTGAGACAGCAA 1014

Qy        380 ArgAlaLeu---GlnGlnIleGlyMetMetGlnProAsnAlaTrpArgProGlnArgGly 398
              |||||||||   ||||||:::||||||::::::|||   ||||||||||||||||||||
Db       1015 AGAGCTTTGCATCAACAGCTTGGAATGGTTCGACCT---GCTTGGAGACCTCAAAGAGGT 1071

Qy        399 LeuProGluArgAlaValSerValLeuArgAlaTrpLeuPheGluHisPheLeuHisPro 418
              |||||||||   :::||||||||||||||||||||||||||||||||||||||||||||
Db       1072 CTTCCTGAAAACTCTGTCTCTGTTCTTCGCGCTTGGCTCTTTGAACATTTCCTTCATCCA 1131

Qy        419 TyrProLysAspSerAspLysIleMetLeuAlaLysGlnThrGlyLeuThrArgSerGln 438
              |||||||||:::|||:::||||||||||||||||||||||||||||||||:::|||
Db       1132 TATCCGAAAGAATCTGAGAAAATCATGCTTGCAAAGCAGACAGGATTGTCAAAGAACCAG 1191

Qy        439 ValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLysProMetValGluGluMet 458
              |||:::||||||||||||||||||||||||||||||||||||||||:::|||||||||
Db       1192 GTTGCTAATTGGTTCATAAACGCGAGAGTTCGTCTGTGGAAACCGATGATCGAAGAAATG 1251

Qy        459 TyrLeuGluGluVal----------------------------LysAsnGlnGluGln 468
              |||   ||||||                             :::|||||||
Db       1252 TATAAAGAAGAGTTTGGTGATGAATCTGAGTTACTAATCTCCAAATCTTCCCAAGAACCC 1311

Qy        469 AsnSerThrAsnThrSerGlyAspAsn------------------------LysAsnLys 480
              ||||||||||||            :::                       |||
Db       1312 AACAGCACAAACCAAGAAGACTCCTCATCGCAGCAGCAGCAGCAGCAAGAGAACAACAAC 1371

Qy        481 GluThrAsnIleSerAlaProAsnGluGluLysHisProIleIleThrSerSer------ 498
              :::|||::::::      :::   :::          |||:::   ||||||
Db       1372 AACAGCAACCTCGCTTATTCATCTGCAGACACAACAAACATTGTCTTCTCATCAGAAACC 1431

Qy        499 ------------LeuLeuGlnAspGlyIleThrThrThrGlnAlaGluIleSerThrSer 514
                       :::|||   :::                 |||   :::|||:::   |||
Db       1432 AAACCAGATCGTGTTCTAGGCAATGATAACGACCCACAGCAACAACAGATAAACCGCTCA 1491

Qy        515 Thr-----------------------------------------------------Ile 516
              :::                                                     |||
Db       1492 TCGGATTACGACACTCTGATGAACTATCACGGGTTTGGTGTTGATGATTACCGTTACATT 1551

Qy        517 SerThrSer---------ProThrAlaGlyAlaSerLeuHisHisAlaHisAsnPhe 532
              |||   |||           :::     ::::::   ||||||   |||:::|||
Db       1552 AGCGGAAGCAACCAGCAAGAAAGCAGATTCTCCAATTCCCATCACTTACACGACTTT 1608

RESULT 8
ABZ13100
ID   ABZ13100 standard; DNA; 1884 BP.
```

```
XX
AC   ABZ13100;
XX
DT   21-JAN-2003  (first entry)
XX
DE   Arabidopsis thaliana stress regulated gene SEQ ID NO 905.
XX
KW   Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX
OS   Arabidopsis thaliana.
XX
PN   WO200216655-A2.
XX
PD   28-FEB-2002.
XX
PF   24-AUG-2001; 2001WO-US026685.
XX
PR   24-AUG-2000; 2000US-0227866P.
PR   26-JAN-2001; 2001US-0264647P.
PR   22-JUN-2001; 2001US-0300111P.
XX
PA   (SCRI ) SCRIPPS RES INST.
PA   (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI   Harper JF,  Kreps J,  Wang X,  Zhu T;
XX
DR   WPI; 2002-304127/34.
XX
PT   Identifying a stress condition to which a plant cell has been exposed and
PT   producing plants with increased tolerance to these abiotic stresses.
XX
PS   Claim 144; SEQ ID NO 905; 577pp + Sequence Listing; English.
XX
CC   The invention relates to identifying a stress condition to which a plant
CC   cell has been exposed, comprising: (a) contacting nucleic acid
CC   representative of expressed polynucleotides in the plant cell with an
CC   array or probes representative of the plant cell genome; and (b)
CC   detecting a profile of expressed polynucleotides in the plant cell
CC   characteristic of a stress response. The method is useful in the
CC   production of transgenic plants, cells and seeds and in producing plants
CC   with increased tolerance to abiotic stress. The present sequence is that
CC   of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
CC   in methods of the invention. Note: The sequence data for this patent is
CC   not represented in the printed specification but is based on sequence
CC   information supplied to Derwent by the European Patent Office
XX
SQ   Sequence 1884 BP; 555 A; 448 C; 418 G; 463 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.6e-58       Length:          1884
Score:                  824.00        Matches:         237
Percent Similarity:     45.96%        Conservative:    76
Best Local Similarity:  34.80%        Mismatches:      178
Query Match:            23.13%        Indels:          190
DB:                     6             Gaps:            25

US-10-624-201A-2 (1-688) x ABZ13100 (1-1884)

Qy         15 AspHisGlnGlnArgHisAsnHisGlySerAsnAsnAsnAsnIleGlnThrLeuTyr 34
              :::   |||   ||||||               ||||||||||||      ||||||::: 
Db        259 GAGCAACAAAACCGTCAC--------------AACAACAACAATAACTCAACTCTTCAT 303

Qy         35 LeuMetAsnProAsnAsnTyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGlnGln 54
              ::::::   ||||||:::::   ||||||:::      :::|||             |||
Db        304 ATGTTATTACCAAATCATCAT---CAAGGTTTTGCTTTCACCGACGAAAACACTATGCAG 360

Qy         55 GlnLeuLeuPheLeuAsnSerSerProAlaAlaSerAsnAlaLeuCysHisAlaAsnIle 74

Db        360 ------------------------------------------------------------ 360

Qy         75 GlnHisAlaProLeuGlnGlnGlnHisPheValGlyValProLeuProAlaValSerLeu 94
                 |||   ||||||||||||||||              |||:::   |||
Db        361 ---------CCGCAGCAACAACAACACTTTACA---------TGGCCATCTTCCTCC--- 399

Qy         95 HisAspGlnIleAsnHisHisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnSer 114
                      :::||||||                        |||:::|||
Db        400 ---------TCCGATCATCAT--------------------CAAAACCGAGATATGATC 429

Qy        115 GlnGlnValIleValProSerSerThrGlyValSerAlaThrSerCysGlyGlyIleThr 134
                 |||   |||             |||:::|||   :::            ::::::
Db        430 GGAACCGTCCACGTGGAAGGAGGAAAGGGTTTGTCTTTATCT-----------CTCTCA 477

Qy        135 ThrAspLeuAlaSerGlnLeuAlaPheGln---ArgProIle----------------- 147
                 :::   ||||||:::       |||   :::   |||   |||
Db        478 TCTTCATTAGCCGCAGCTAAAGCCGAGGAATATAGAAGCATTTATTGTGCAGCCGTTGAT 537

Qy        148 ------------------ProThrProGlnHisArgGlnGlnGlnGlnGlnGlyGly 161
                                     |||       |||       |||
Db        538 GGAACTTCTTCTTCTTCTAACGCATCCGCTCATCATCATCAATTCAATCAGTTCAAGAAT 597

Qy        162 LeuSerLeu------SerLeuSerProGlnLeuGlnGlnGlnIleSerPheAsnAsnAsn 179
                 |||   |||      |||   |||                 |||:::      :::
Db        598 CTTCTTCTTGAGAATTCTTCTTCTCAACATCATCACCATCAAGTTGTTGGACATTTTGGT 657

Qy        180 IleSerSerSerSerProArgThrAsnAsnValThrIleArgGlyThrLeuAspGlySer 199
                    ||||||||||||||         :::   :::|||   |||
Db        658 TCATCATCATCATCTCCCATGGCGGCTTCTTCATCCATTGGAGGGATCTAC--------- 708

Qy        200 SerSerAsnMetValLeuGlySerLysTyrLeuLysAlaAlaGlnGluLeuLeuAspGlu 219
                          :::      |||||||||   |||   |||||||||||||||:::|||
Db        709 ---------ACGTTGAGGAATTCGAAATATACGAAACCGGCTCAAGAGTTGTTGGAAGAG 759

Qy        220 ValValAsnIleValGlyLysSerIleLysGlyAspAspGlnLysLysAspAsnSerMet 239
                 ::::::                  |||           |||   |||   :::
Db        760 TTTTGTAGTGTT-----------------GGAAGAGGACATTTCAAGAAGAACAAACTT 801

Qy        240 AsnLysGluSerMetProLeuAlaSerAspValAsnThr------AsnSerSerGlyGly 257
                 ::::::   :::            |||:::   ||||||                ||||||
Db        802 AGTAGGAACAAC------------TCAAACCCTAATACTACCGGTGGAGGAGGAGGCGGA 849

Qy        258 GlyGluSerSerSerArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGlnArg 277
                 |||   |||||||||              :::      |||:::   |||   |||
Db        850 GGGTCCTCGTCATCGGCCGGAACAGCTAATGATAGTCCTCCTTTGTCTCCGGCTGATCGG 909

Qy        278 GlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeuGluGluValGluGlnArg 297
                    |||   |||   :::|||   |||||||||:::|||||||||||||||:::::|||
Db        910 ATTGAACATCAAAGAAGAAAAGTCAAGCTACTATCTATGCTTGAAGAGGTGGACCGACGG 969

Qy        298 TyrArgGlnTyrHisHisGlnMetGlnIleIleValLeuSerPheGluGlnValAlaGly 317
                 |||       |||       |||||||||:::::|||   ||||||:::||||||   |||
Db        970 TACAACCACTACTGCGAACAAATGCAAATGGTAGTGAACTCATTCGACCAAGTAATGGGT 1029

Qy        318 IleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArg 337
                    |||:::|||       ||||||   ||||||       |||:::|||:::   ||||||
Db       1030 TACGGCGCGGCGGTTCCGTACACGACATTAGCTCAAAAGGCAATGTCTAGGCATTTCCGG 1089

Qy        338 CysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGluGlu 357
                 ||||||||||||||:::|||   |||:::|||   :::   :::   ||||||::::::
Db       1090 TGTTTGAAAGACGCGGTAGCGGTTCAGCTTAAACGCAGCTGTGAGCTTCTAGGGGATAAA 1149

Qy        358 GluGlyLeuGlyGly-----------------LysIleGluGlySerArgLeuLysPhe 371
                 |||       |||               |||   |||       ||||||:::
Db       1150 GAGGCGGCAGGGGCTGCATCCTCGGGGTTAACCAAAGGGGAAACGCCGCGATTGCGTTTG 1209

Qy        372 ValAspHisHisLeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGlnProAsn 391
                 ::::::      ||||||||||||||||||         :::||||||||:::
Db       1210 CTAGAGCAGAGTTTGCGTCAGCAACGAGCGTTTCATCATATGGGTATGATGGAGCAAGAG 1269

Qy        392 AlaTrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeu 411
                 ||||||||||||||||||||||||||||||:::|||::::::|||||||||||||||
Db       1270 GCATGGAGACCGCAACGTGGTTTGCCTGAACGCTCCGTTAATATCCTTAGAGCTTGGCTA 1329

Qy        412 PheGluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGln 431
                 ||||||||||||||:::||||||||||   |||:::||||||   :::|||||:::|||
Db       1330 TTCGAGCATTTTCTTAATCCGTACCCAAGCGATGCTGATAAGCACCTCTTAGCACGACAG 1389
```

```
Qy        432 ThrGlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrp 451
              ||||||||||:::|||:::||||||||||||||||||||||||||||||||||||||||||
Db       1390 ACTGGTTTATCCAGAAATCAGGTGTCAAATTGGTTCATAAATGCTAGGGTTCGCCTATGG 1449

Qy        452 LysProMetValGluGluMetTyrLeuGluGluValLys---------------------- 464
              ||||||||||||||||||||||||   :::|||   |||
Db       1450 AAACCAATGGTGGAAGAGATGTATCAACAAGAAGCAAAAGAAAGAGAAGAAGCAGAAGAA 1509

Qy        465 ------------AsnGlnGluGlnAsnSerThrAsnThrSerGlyAspAsnLysAsnLys 480
                          |||:::|||                :::   :::|||      |||
Db       1510 GAAAATGAAAATCAACAACAACAAAGAAGACAGCAACAAACAAACAACAACGACACGAAA 1569

Qy        481 GluThrAsnIleSerAlaProAsnGluGluLysHisProIleIleThrSerSerLeuLeu 500
                 |||          ||||||       :::||||||:::
Db       1570 CCCAACAAC------------AATGAAAACAACTTCACTGTCATAACCGCACAA------ 1611

Qy        501 GlnAspGlyIleThrThrThrGlnAlaGluIleSerThrSerThrIleSerThrSerPro 520
                                                                   :::|||
Db       1612 ----------------------------------------------------ACTCCA 1617

Qy        521 ThrAlaGlyAlaSerLeuHisHisAlaHisAsnPheSerPheLeuGlySerPheAsnMet 540
               |||          |||   ||||||
Db       1618 ACGACGATGACATCGACACATCAC---------------------------------- 1641

Qy        541 AspAsnThrThrThrThrValAspHisIleGluAsnAsnAlaLysLysGlnArgAsnAsp 560
                                                                  ||||||
Db       1642 ----------------------------------------------------GAAAACGAC 1650

Qy        561 MetHisLysPheSerProSerSerIleLeuSerSerValAspMetGluAlaLysAlaArg 580
                          ||||||   ||||||||||||          |||   |||
Db       1651 -----------------TCTTCATTCCTCTCTTCCGTC---------GCCGCCGCTTCT 1683

Qy        581 GluSerSerAsnLysGlyPheThrAsnProLeuMetAlaAlaTyrAlaMetGlyAspPhe 600
                     :::     ||||||
Db       1684 CACGGCGGTTCAGACGCGTTCACCGTCGCCACGTGT----------------------- 1719

Qy        601 GlyArgPheAspProHisAspGlnGlnMetThrAlaAsnPheHis------------Gly 616
                          ||||||     ::::::||||||             |||
Db       1720 ----------------------CAGCAAGACGTCAGTGACTTCCACGTCGACGGAGATGGT 1758

Qy        617 AsnAsn-----------------------------GlyValSerLeuThrLeuGlyLeu 626
                 |||                           ||||||||||||||||||||
Db       1759 GTGAACGTCATAAGATTCGGGACCAAACAGACTGGTGACGTGTCTCTTACGCTTGGTCTA 1818

Qy        627 ProProSerGluAsnLeuAlaMetProValSerGlnGlnAsnTyrLeuSerAsnAspLeu 646
                  |||   |||:::      |||            ::::::           |||
Db       1819 CGCCACTCTGGCAATATT------CCTGATAAGAACACTTCTTTCTCCGTTAGAGACTTT 1872

Qy        647 Gly 647
                 |||
Db       1873 GGA 1875
```

```
RESULT 9
AAD06494
ID   AAD06494 standard; cDNA; 1983 BP.
XX
AC   AAD06494;
XX
DT   10-AUG-2001  (first entry)
XX
DE   Arabidopsis thaliana G418 transcription factor homologue, G965 cDNA.
XX
KW   Transcription factor; pesticidal; antimicrobial; gene therapy;
KW   pathogen tolerance; trichome structure; callose induction;
KW   phytoalexin induction; plant structure; plant development; ss.
XX
OS   Arabidopsis thaliana.
XX
FH   Key             Location/Qualifiers
FT   CDS             73. .1956
PT                   /*tag=  a
FT                   /product= "A. thaliana transcription factor homologue"
XX
PN   WO200135726-A1.
XX
PD   25-MAY-2001.
XX
PF   14-NOV-2000; 2000WO-US031418.
XX
PR   17-NOV-1999;   99US-0166228P.
PR   17-APR-2000; 2000US-0197899P.
PR   22-AUG-2000; 2000US-0227439P.
XX
PA   (MEND-) MENDEL BIOTECHNOLOGY INC.
PA   (HEAR/) HEARD J.
PA   (RATC/) RATCLIFFE O.
PA   (CREE/) CREELMAN R.
PA   (JIAN/) JIANG C.
PA   (PINE/) PINEDA O.
PA   (REUB/) REUBER L.
PA   (ADAM/) ADAM L.
XX
PI   Heard J,  Ratcliffe O,  Creelman R,  Jiang C,  Pineda O,  Reuber L;
PI   Adam L;
XX
DR   WPI; 2001-335978/35.
DR   P-PSDB; AAE02525.
XX
PT   Nucleic acids encoding plant transcription factor polypeptides, useful
PT   for altering the pathogen resistance characteristics of plants, e.g.
PT   corn, potato and cotton plants.
XX
PS   Claim 4; Page 114-116; 134pp; English.
XX
CC   The present sequence is a cDNA encoding Arabidopsis thaliana
CC   transcription factor homologue. The transcription factors are used to
CC   alter the structure and developmental characteristics of plants such as
CC   soybean, wheat, corn, potato, cotton, rice, oilseed rape, sunflower,
CC   alfalfa, sugarcane, turf, banana, blackberry, blueberry, strawberry,
CC   raspberry, carrot, cantaloupe, cauliflower, cucumber, coffee, eggplant,
CC   grapes, mango, lettuce, honeydew, melon, onion, papaya, peppers,
CC   pineapple, spinach, squash, sweet corn, tobacco, tomato, peas,
CC   watermelon, rosaceous fruits and vegetable brassicas. The transcription
CC   factors are specifically useful for modifying traits associated with
CC   plant's pathogen tolerance such as alterations in cell wall composition,
CC   trichome number or structure, callose induction, phytoalexin induction,
CC   and alterations in the cell death response. Transgenic plants expressing
CC   these transcription factors are more tolerant to biotrophic or
CC   necrotrophic pathogens such as fungi, bacteria, mollicutes, viruses,
CC   nematodes and parasitic higher plants. The transcription factors are also
CC   used in gene therapy
XX
SQ   Sequence 1983 BP; 585 A; 464 C; 434 G; 500 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.:                1.72e-58        Length:          1983
Score:                    824.00          Matches:         237
Percent Similarity:       45.96%          Conservative:    76
Best Local Similarity:    34.80%          Mismatches:      178
Query Match:              23.13%          Indels:          190
DB:                       4               Gaps:            25

US-10-624-201A-2 (1-688) x AAD06494 (1-1983)

Qy         15 AspHisGlnGlnArgHisAsnHisGlyAsnSerAsnAsnAsnAsnIleGlnThrLeuTyr 34
              :::   |||   ||||||               ||||||||||||      ||||||:::
Db        331 GAGCAACAAAAACCGTCAC--------------AACAACAACAATAACTCAACTCTTCAT 375

Qy         35 LeuMetAsnProAsnAsnTyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGlnGln 54
              ::::::   |||||::::::   ||||||:::      :::|||            |||
Db        376 ATGTTATTACCAAATCATCAT---CAAGGTTTTGCTTTCACCGACGAAAACACTATGCAG 432

Qy         55 GlnLeuLeuPheLeuAsnSerSerProAlaAlaSerAsnAlaLeuCysHisAlaAsnIle 74

Db        432 ---------------------------------------------------------- 432
```

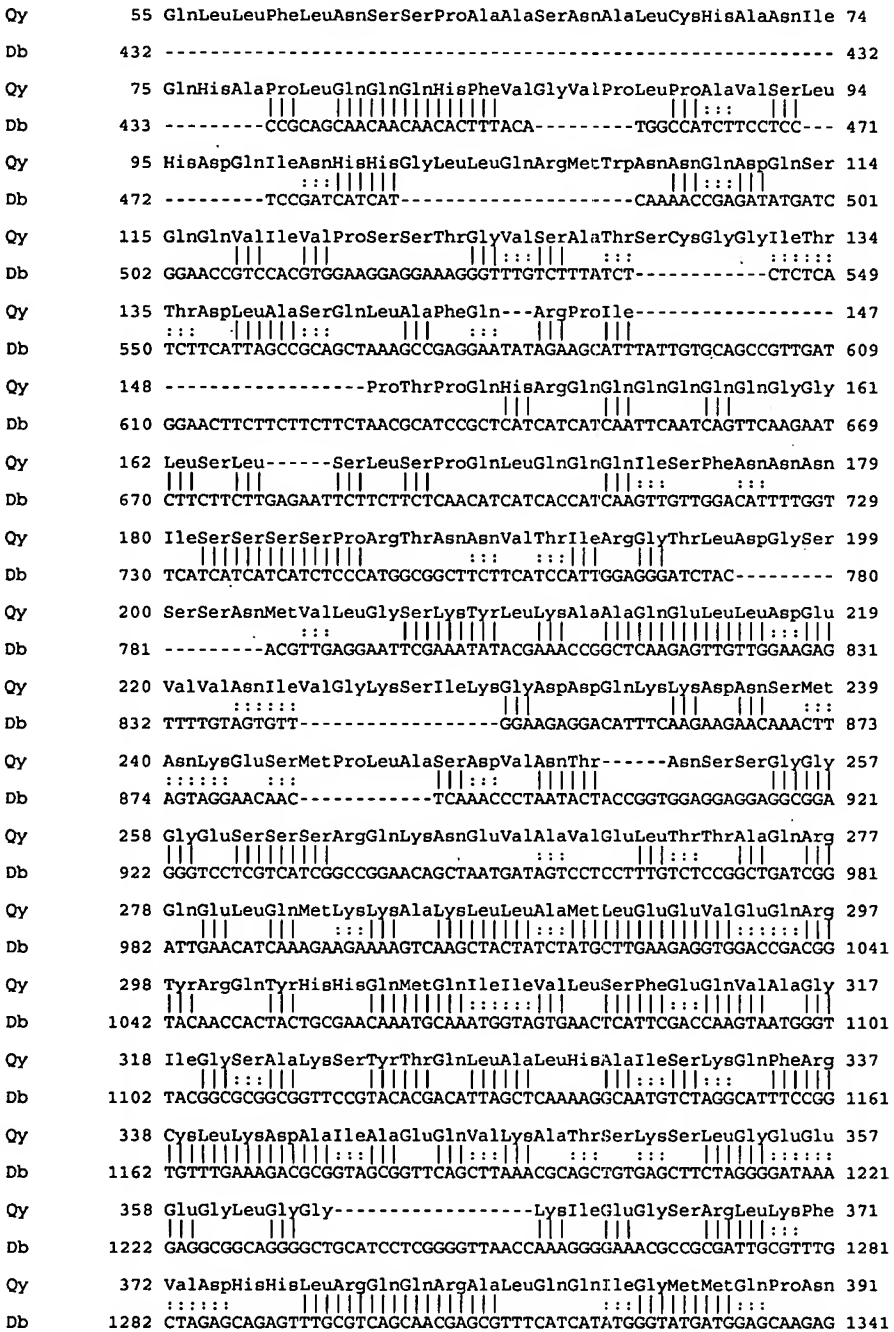


RESULT 10
ADC46618
ID ADC46618 standard; DNA; 1983 BP.
XX
AC ADC46618;
XX
DT 18-DEC-2003 (first entry)
XX
DE Thalecress transcription factor-like DNA G965.
XX
KW Thalecress; transcription factor-like protein; ds; seed trait;
KW transgenic; gene; plant size; stress tolerance; yield;
KW disease resistance; plant.
XX
OS Arabidopsis thaliana.
XX
PN US2003093837-A1.
XX

```
PD   15-MAY-2003.
XX
PF   01-NOV-2002; 2002US-00286264.
XX
PR   23-MAR-1999;   99US-0125814P.
PR   22-MAR-2000; 2000US-00533030.
XX
PA   (KEDD/) KEDDIE J.
PA   (RIEC/) RIECHMANN J L.
PA   (RATC/) RATCLIFFE O.
PA   (ZHAN/) ZHANG J.
PA   (JIAN/) JIANG C.
PA   (PINE/) PINEDA O.
PA   (HEAR/) HEARD J.
PA   (YUGG/) YU G.
PA   (ADAM/) ADAM L.
PA   (BROU/) BROUN P.
PA   (REUB/) REUBER L.
PA   (PILG/) PILGRIM M.
PA   (SAMA/) SAMAHA R.
XX
PI   Keddie J,  Riechmann JL,  Ratcliffe O,  Zhang J,  Jiang C,  Pineda O;
PI   Heard J,  Yu G,  Adam L,  Broun P,  Reuber L,  Pilgrim M,  Samaha R;
XX
DR   WPI; 2003-765498/72.
DR   P-PSDB; ADC46619.
XX
PT   Novel transgenic plant having recombinant polynucleotide encoding
PT   polypeptide that alters trait of transgenic plant when compared with same
PT   trait of another plant lacking recombinant polynucleotide.
XX
PS   Disclosure; SEQ ID NO 17; 165pp; English.
XX
CC   The invention relates to a transgenic plant having recombinant
CC   polynucleotide (II) encoding polypeptide comprising at least 6
CC   consecutive amino acids of a sequence chosen from the protein sequence
CC   appearing as ADC46603 - ADC46749 (every second sequence), where
CC   recombinant polynucleotide alters a trait of the seed transgenic plant
CC   when compared with same trait of another plant lacking recombinant
CC   polynucleotide. The proteins are transcription factor-like proteins. Also
CC   included are altering (M1) a trait associated with seed (comprising:
CC   transforming a plant with (II); selecting the transformed plants; and
CC   identifying a transformed plant with seed having altered trait), altering
CC   (M2) the expression levels of at least one gene of a plant (involving
CC   transforming the plant with (II) and selecting the transformed plant),
CC   altering (M3) a trait associated with a plant's seed (comprising:
CC   transforming the plant with a recombinant polynucleotide comprising a
CC   nucleotide sequence comprising least 18 consecutive nucleotides of a
CC   sequence appearing as ADC46750 - ADC46766 and selecting the transformed
CC   plan) altering (M4) a plant's trait (involving providing a database
CC   sequence, comparing the database sequence with a polypeptide or a
CC   polynucleotide chosen as detailed above, selecting a database sequence
CC   that needs selected sequence criteria and transforming a database
CC   sequence in the plant) and altering a plant's trait (involving providing
CC   a test polynucleotide, hybridising the test polynucleotide with a
CC   polynucleotide detailed above at low stringency and transforming the
CC   hybridising test polynucleotide in a plant to alter a trait of the
CC   plant). The method (M1) is useful for altering a trait associated with
CC   seed. The method (M2) is useful for altering the expressing levels of at
CC   least one gene of a plant. The method (M3) is useful for altering a trait
CC   associated with a plant's seed. The method (M4) is useful for altering a
CC   plant's trait . The method (M4) is useful for altering a plant's trait
CC   such as seed or plant size, stress tolerance, yield or disease
CC   resistance. The present sequence encodes a transcription factor-like
CC   protein/seed trait altering protein of the invention.
XX
SQ   Sequence 1983 BP; 586 A; 464 C; 433 G; 500 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.72e-58        Length:         1983
Score:                  824.00          Matches:        237
Percent Similarity:     45.96%          Conservative:   76
Best Local Similarity:  34.80%          Mismatches:     178

Query Match:            23.13%          Indels:         190
DB:                     10              Gaps:           25

US-10-624-201A-2 (1-688) x ADC46618 (1-1983)

Qy         15 AspHisGlnGlnArgHisAsnHisGlyAsnSerAsnAsnAsnAsnIleGlnThrLeuTyr 34
Db        331 GAGCAACAAAACCGTCAC--------------AACAACAACAATAACTCAACTCTTCAT 375

Qy         35 LeuMetAsnProAsnAsnTyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGlnGln 54
Db        376 ATGTTATTACCAAATCATCAT---CAAGGTTTTGCTTTCACCGACGAAAACACTATGCAG 432

Qy         55 GlnLeuLeuPheLeuAsnSerSerProAlaAlaSerAsnAlaLeuCysHisAlaAsnIle 74
Db        432 ------------------------------------------------------------ 432

Qy         75 GlnHisAlaProLeuGlnGlnGlnHisPheValGlyValProLeuProAlaValSerLeu 94
Db        433 ---------CCGCAGCAACAACAACACTTTACA---------TGGCCATCTTCCTCC--- 471

Qy         95 HisAspGlnIleAsnHisHisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnSer 114
Db        472 ---------TCCGATCATCAT--------------------CAAAACCGAGATATGATC 501

Qy        115 GlnGlnValIleValProSerSerThrGlyValSerAlaThrSerCysGlyGlyIleThr 134
Db        502 GGAACCGTCCACGTGGAAGGAGGAAAGGGTTTGTCTTTATCT-----------CTCTCA 549

Qy        135 ThrAspLeuAlaSerGlnLeuAlaPheGln---ArgProIle----------------- 147
Db        550 TCTTCATTAGCCGCAGCTAAAGCCGAGGAATATAGAAGCATTTATTGTGCAGCCGTTGAT 609

Qy        148 ------------------ProThrProGlnHisArgGlnGlnGlnGlnGlnGlyGly 161
Db        610 GGAACTTCTTCTTCTTCTAACGCATCCGCTCATCATCATCAATTCAATCAGTTCAAGAAT 669

Qy        162 LeuSerLeu------SerLeuSerProGlnLeuGlnGlnGlnIleSerPheAsnAsnAsn 179
Db        670 CTTCTTCTTGAGAATTCTTCTTCTCAACATCATCACCATCAAGTTGTTGGACATTTTGGT 729

Qy        180 IleSerSerSerSerProArgThrAsnAsnValThrIleArgGlyThrLeuAspGlySer 199
Db        730 TCATCATCATCATCTCCCATGGCGGCTTCTTCATCCATTGGAGGGATCTAC--------- 780

Qy        200 SerSerAsnMetValLeuGlySerLysTyrLeuLysAlaAlaGlnGluLeuLeuAspGlu 219
Db        781 ---------ACGTTGAGGAATTCGAAATATACGAAACCGGCTCAAGAGTTGTTGGAAGAG 831

Qy        220 ValValAsnIleValGlyLysSerIleLysGlyAspAspGlnLysLysAspAsnSerMet 239
Db        832 TTTTGTAGTGTT-----------------GGAAGAGGACATTTCAAGAAGAACAAACTT 873

Qy        240 AsnLysGluSerMetProLeuAlaSerAspValAsnThr------AsnSerSerGlyGly 257
Db        874 AGTAGGAACAAC------------TCAAACCCTAATACTACCGGTGGAGGAGGAGGCGGA 921

Qy        258 GlyGluSerSerSerArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGlnArg 277
Db        922 GGGTCCTCGTCATCGGCCGGAACAGCTAATGATAGTCCTCCTTTGTCTCCGGCTGATCGG 981

Qy        278 GlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeuGluGluValGluGlnArg 297
Db        982 ATTGAACATCAAAGAAGAAAAGTCAAGCTACTATCTATGCTTGAAGAGGTGGACCGACGG 1041

Qy        298 TyrArgGlnTyrHisHisGlnMetGlnIleIleValLeuSerPheGluGlnValAlaGly 317
Db       1042 TACAACCACTACTGCGAACAAATGCAAATGGTAGTGAACTCATTCGACCAAGTAATGGGT 1101

Qy        318 IleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArg 337
Db       1102 TACGGCGCGGCGGTTCCGTACACGACATTAGCTCAAAAGGCAATGTCTAGGCATTTCCGG 1161
```

```
Qy        338 CysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGluGlu 357
              ||||||||||||||||:::|||   |||:::|||   :::   :::   ||||||:::::: 
Db       1162 TGTTTGAAAGACGCGGTAGCGGTTCAGCTTAAACGCAGCTGTGAGCTTCTAGGGGATAAA 1221

Qy        358 GluGlyLeuGlyGly-----------------LysIleGluGlySerArgLeuLysPhe 371
              |||      |||                    |||   |||      ||||||::: 
Db       1222 GAGGCGGCAGGGGCTGCATCCTCGGGGTTAACCAAAGGGGAAACGCCGCGATTGCGTTTG 1281

Qy        372 ValAspHisHisLeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGlnProAsn 391
              ::::::      ||||||||||||||||||         :::|||||||||:::
Db       1282 CTAGAGCAGAGTTTGCGTCAGCAACGAGCGTTTCATCATATGGGTATGATGGAGCAAGAG 1341

Qy        392 AlaTrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeu 411
              ||||||||||||||||||||||||||||||:::|||::::::|||||||||||||||| 
Db       1342 GCATGGAGACCGCAACGTGGTTTGCCTGAACGCTCCGTTAATATCCTTAGAGCTTGGCTA 1401

Qy        412 PheGluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGln 431
              ||||||||||||||:::|||||||||   |||:::||||||   :::||||||:::||| 
Db       1402 TTCGAGCATTTTCTTAATCCGTACCCAAGCGATGCTGATAAGCACCTCTTAGCACGACAG 1461

Qy        432 ThrGlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrp 451
              |||||||||:::|||:::|||||||||||||||||||||||||||||||||||||||||| 
Db       1462 ACTGGTTTATCCAGAAATCAGGTGTCAAATTGGTTCATAAATGCTAGGGTTCGCCTATGG 1521

Qy        452 LysProMetValGluGluMetTyrLeuGluGluValLys---------------------- 464
              ||||||||||||||||||||||   :::|||   |||
Db       1522 AAACCAATGGTGGAAGAGATGTATCAACAAGAAGCAAAAGAAAGAGAAGAAGCAGAAGAA 1581

Qy        465 ------------AsnGlnGluGlnAsnSerThrAsnThrSerGlyAspAsnLysAsnLys 480
                         |||:::|||                :::   :::|||      |||
Db       1582 GAAAATGAAAATCAACAACAACAAAGAAGACAGCAACAAACAAACAACAACGACACGAAA 1641

Qy        481 GluThrAsnIleSerAlaProAsnGluGluLysHisProIleIleThrSerSerLeuLeu 500
              |||            ||||||          :::|||||:::
Db       1642 CCCAACAAC------------AATGAAAACAACTTCACTGTCATAACCGCACAA------ 1683

Qy        501 GlnAspGlyIleThrThrThrGlnAlaGluIleSerThrSerThrIleSerThrSerPro 520
                                                                     :::|||
Db       1684 ------------------------------------------------------ACTCCA 1689

Qy        521 ThrAlaGlyAlaSerLeuHisHisAlaHisAsnPheSerPheLeuGlySerPheAsnMet 540
              |||          |||   ||||||
Db       1690 ACGACGATGACATCGACACATCAC---------------------------------- 1713

Qy        541 AspAsnThrThrThrThrValAspHisIleGluAsnAsnAlaLysLysGlnArgAsnAsp 560
                                                                   ||||||
Db       1714 ------------------------------------------------GAAAACGAC 1722

Qy        561 MetHisLysPheSerProSerSerIleLeuSerSerValAspMetGluAlaLysAlaArg 580
                          ||||||   ||||||||||||         |||   |||
Db       1723 -----------------TCTTCATTCCTCTCTTCCGTC---------GCCGCCGCTTCT 1755

Qy        581 GluSerSerAsnLysGlyPheThrAsnProLeuMetAlaAlaTyrAlaMetGlyAspPhe 600
                    :::     ||||||
Db       1756 CACGGCGGTTCAGACGCGTTCACCGTCGCCACGTGT----------------------- 1791

Qy        601 GlyArgPheAspProHisAspGlnGlnMetThrAlaAsnPheHis-----------Gly 616
                               ||||||      ::::::||||||             |||
Db       1792 --------------------CAGCAAGACGTCAGTGACTTCCACGTCGACGGAGATGGT 1830

Qy        617 AsnAsn----------------------------GlyValSerLeuThrLeuGlyLeu 626
                 |||                           ||||||||||||||||||||
Db       1831 GTGAACGTCATAAGATTCGGGACCAAACAGACTGGTGACGTGTCTCTTACGCTTGGTCTA 1890

Qy        627 ProProSerGluAsnLeuAlaMetProValSerGlnGlnAsnTyrLeuSerAsnAspLeu 646
                    |||   |||:::      |||           ::::::         |||
Db       1891 CGCCACTCTGGCAATATT------CCTGATAAGAACACTTCTTTCTCCGTTAGAGACTTT 1944

Qy        647 Gly 647
              |||
Db       1945 GGA 1947

RESULT 11
```

```
ADD30786
ID   ADD30786 standard; cDNA; 1983 BP.
XX
AC   ADD30786;
XX
DT   15-JAN-2004  (first entry)
XX
DE   Plant yield-related polynucleotide clone G965.
XX
KW   ds; transcription factor; transgenic plant; growth rate; senescence;
KW   seed germination rate; plant vigor; seedling vigor.
XX
OS   Arabidopsis thaliana.
XX
PN   WO2003013227-A2.
XX
PD   20-FEB-2003.
XX
PF   09-AUG-2002; 2002WO-US025805.
XX
PR   09-AUG-2001; 2001US-0310847P.
PR   19-NOV-2001; 2001US-0336049P.
PR   11-DEC-2001; 2001US-0338692P.
PR   14-JUN-2002; 2002US-00171468.
XX
PA   (MEND-) MENDEL BIOTECHNOLOGY INC.
XX
PI   Ratcliffe O,  Riechmann JL,  Adam LJ,  Dubell AT,  Heard JE;
PI   Pilgrim ML,  Jiang C,  Reuber TL,  Creelman RA,  Pineda O,  Yu G;
PI   Broun PE;
XX
DR   WPI; 2003-248221/24.
DR   P-PSDB; ADD30787.
XX
PT   New plant transcription factor polynucleotides and polypeptides, useful
PT   in producing transgenic plants with commercially valuable properties,
PT   such as an alteration in a plant growth characteristic, e.g. growth rate
PT   or apomixis.
XX
PS   Disclosure; SEQ ID NO 815; 454pp; English.
XX
CC   The invention relates to a number of isolated Arabidopsis thaliana cDNA
CC   sequences and their encoded proteins which are especially transcription
CC   factor related cDNA's and proteins. The isolated or recombinant plant
CC   transcription factor polynucleotides and polypeptides are useful in
CC   producing transgenic plants with commercially valuable properties, i.e.
CC   modified or altered desirable traits as compared to a reference plant,
CC   such as an alteration in a plant growth characteristic, e.g. growth rate,
CC   germination rate of seeds, vigor of plants and seedlings, or leaf and
CC   flower senescence. Sequence information related to the polynucleotides
CC   and polypeptides can also be used in bioinformatic search methods. The
CC   transgenic plant is useful for growing a progeny plant from a parent
CC   plant. This sequence represents one of the cDNAs of the invention.
XX
SQ   Sequence 1983 BP; 585 A; 464 C; 434 G; 500 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.72e-58        Length:         1983
Score:                  824.00          Matches:        237
Percent Similarity:     45.96%          Conservative:   76
Best Local Similarity:  34.80%          Mismatches:     178
Query Match:            23.13%          Indels:         190
DB:                     10              Gaps:           25

US-10-624-201A-2 (1-688) x ADD30786 (1-1983)

Qy         15 AspHisGlnGlnArgHisAsnHisGlyAsnSerAsnAsnAsnAsnIleGlnThrLeuTyr 34
                 :::   |||   ||||||               ||||||||||||      ||||||::: 
Db        331 GAGCAACAAAACCGTCAC--------------AACAACAACAATAACTCAACTCTTCAT 375

Qy         35 LeuMetAsnProAsnAsnTyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGlnGln 54
                 ::::::   ||||||::::::   ||||||:::      :::|||            |||
Db        376 ATGTTATTACCAAATCATCAT---CAAGGTTTTGCTTTCACCGACGAAAACACTATGCAG 432
```



RESULT 12
ADI44018
ID ADI44018 standard; DNA; 1983 BP.
XX
AC ADI44018;
XX
DT 22-APR-2004 (first entry)
XX
DE Plant transcription factor related polynucleotide #1598.
XX
KW transgenic; plant; enhanced tolerance to abiotic stress;
KW glyphosphate tolerance; hormone sensitivity; disease resistance;
KW sugar sensing; flowering; flower structure; stem bifurcation;

```
KW   branching pattern; apical dominance; trichome; stem morphology;
KW   root growth; root hair; seed development; cell proliferation;
KW   cell differentiation; premature senescence; necrosis; plant size;
KW   leaf morphology; seed morphology; seed biochemistry; root anthocyanin;
KW   plant anthocyanin; light response; shade avoidance; bioinformatic;
KW   transcription factor; ds.
XX
OS   Unidentified.
XX
PN   US2004019927-A1.
XX
PD   29-JAN-2004.
XX
PF   25-FEB-2003; 2003US-00374780.
XX
PR   18-APR-2001; 2001US-00837944.
XX
PA   (SHER/) SHERMAN B K.
PA   (RIEC/) RIECHMANN J L.
PA   (JIAN/) JIANG C.
PA   (HEAR/) HEARD J E.
PA   (HAAK/) HAAKE V.
PA   (CREE/) CREELMAN R A.
PA   (RATC/) RATCLIFFE O.
PA   (ADAM/) ADAM L J.
PA   (REUB/) REUBER T L.
PA   (KEDD/) KEDDIE J.
PA   (BROU/) BROUN P E.
PA   (PILG/) PILGRIM M L.
PA   (DUBE/) DUBELL A N.
PA   (PINE/) PINEDA O.
PA   (YUGG/) YU G.
XX
PI   Sherman BK,  Riechmann JL,  Jiang C,  Heard JE,  Haake V;
PI   Creelman RA,  Ratcliffe O,  Adam LJ,  Reuber TL,  Keddie J,  Broun PE;
PI   Pilgrim ML,  Dubell AN,  Pineda O,  Yu G;
XX
DR   WPI; 2004-132245/13.
DR   P-PSDB; ADI44019.
XX
PT   New transgenic plant comprising a recombinant polynucleotide of any one
PT   of more than 500 nucleotide sequences, useful in bioinformatic search
PT   methods.
XX
PS   Disclosure; SEQ ID NO 2481; 435pp; English.
XX
CC   The invention describes a transgenic plant comprising a recombinant
CC   polynucleotide of any one of more than 500 nucleotide sequences fully
CC   defined in the specification or its complement. The method of the
CC   invention can be used to produced a plant having altered traits such as:
CC   enhanced tolerance to abiotic stress; glyphosphate tolerance; hormone
CC   sensitivity; disease resistance; sugar sensing; early or late flowering;
CC   altered flower structure, change in stem bifurcations, altered branching
CC   pattern, reduced apical dominance, reduced trichome density; lack of
CC   trichomes; reduced ectopic trichome development; altered trichome
CC   development; increase in trichome number; altered stem morphology;
CC   increased root growth; increased root hairs; altered seed development;
CC   altered cell proliferation or cell differentiation; rapid development;
CC   premature senescence; increased necrosis; increase in seedling or plant
CC   size; decreased plant size; leaf morphology; seed morphology; seed
CC   biochemistry; increase in root anthocyanins; increase in plant
CC   anthocyanins, or alteration in light response or shade avoidance. The
CC   transgenic plant, polynucleotides and polypeptides are useful in
CC   bioinformatic search methods. This sequence represents a plant
CC   transcription factor related polynucleotide.
XX
SQ   Sequence 1983 BP; 585 A; 464 C; 434 G; 500 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.72e-58        Length:         1983
Score:                  824.00          Matches:        237
Percent Similarity:     45.96%          Conservative:   76
Best Local Similarity:  34.80%          Mismatches:     178

Query Match:            23.13%          Indels:         190
DB:                     12              Gaps:           25

US-10-624-201A-2 (1-688) x ADI44018 (1-1983)

Qy         15 AspHisGlnGlnArgHisAsnHisGlyAsnSerAsnAsnAsnAsnIleGlnThrLeuTyr 34
Db        331 GAGCAACAAAACCGTCAC--------------AACAACAACAATAACTCAACTCTTCAT 375

Qy         35 LeuMetAsnProAsnAsnTyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGlnGln 54
Db        376 ATGTTATTACCAAATCATCAT---CAAGGTTTTGCTTTCACCGACGAAAACACTATGCAG 432

Qy         55 GlnLeuLeuPheLeuAsnSerSerProAlaAlaSerAsnAlaLeuCysHisAlaAsnIle 74
Db        432 ------------------------------------------------------------ 432

Qy         75 GlnHisAlaProLeuGlnGlnGlnHisPheValGlyValProLeuProAlaValSerLeu 94
Db        433 ---------CCGCAGCAACAACAACACTTTACA---------TGGCCATCTTCCTCC--- 471

Qy         95 HisAspGlnIleAsnHisHisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnSer 114
Db        472 ---------TCCGATCATCAT--------------------CAAAACCGAGATATGATC 501

Qy        115 GlnGlnValIleValProSerSerThrGlyValSerAlaThrSerCysGlyGlyIleThr 134
Db        502 GGAACCGTCCACGTGGAAGGAGGAAAGGGTTTGTCTTTATCT------------CTCTCA 549

Qy        135 ThrAspLeuAlaSerGlnLeuAlaPheGln---ArgProIle------------------ 147
Db        550 TCTTCATTAGCCGCAGCTAAAGCCGAGGAATATAGAAGCATTTATTGTGCAGCCGTTGAT 609

Qy        148 ------------------ProThrProGlnHisArgGlnGlnGlnGlnGlnGlyGly 161
Db        610 GGAACTTCTTCTTCTTCTAACGCATCCGCTCATCATCATCAATTCAATCAGTTCAAGAAT 669

Qy        162 LeuSerLeu------SerLeuSerProGlnLeuGlnGlnGlnIleSerPheAsnAsnAsn 179
Db        670 CTTCTTCTTGAGAATTCTTCTTCTCAACATCATCACCATCAAGTTGTTGGACATTTTGGT 729

Qy        180 IleSerSerSerSerProArgThrAsnAsnValThrIleArgGlyThrLeuAspGlySer 199
Db        730 TCATCATCATCATCTCCCATGGCGGCTTCTTCATCCATTGGAGGGATCTAC--------- 780

Qy        200 SerSerAsnMetValLeuGlySerLysTyrLeuLysAlaAlaGlnGluLeuLeuAspGlu 219
Db        781 ---------ACGTTGAGGAATTCGAAATATACGAAACCGGCTCAAGAGTTGTTGGAAGAG 831

Qy        220 ValValAsnIleValGlyLysSerIleLysGlyAspAspGlnLysLysAspAsnSerMet 239
Db        832 TTTTGTAGTGTT-----------------GGAAGAGGACATTTCAAGAAGAACAAACTT 873

Qy        240 AsnLysGluSerMetProLeuAlaSerAspValAsnThr------AsnSerSerGlyGly 257
Db        874 AGTAGGAACAAC------------TCAAACCCTAATACTACCGGTGGAGGAGGAGGCGGA 921

Qy        258 GlyGluSerSerSerArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGlnArg 277
Db        922 GGGTCCTCGTCATCGGCCGGAACAGCTAATGATAGTCCTCCTTTGTCTCCGGCTGATCGG 981

Qy        278 GlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeuGluGluValGluGlnArg 297
Db        982 ATTGAACATCAAAGAAGAAAAGTCAAGCTACTATCTATGCTTGAAGAGGTGGACCGACGG 1041

Qy        298 TyrArgGlnTyrHisHisGlnMetGlnIleIleValLeuSerPheGluGlnValAlaGly 317
Db       1042 TACAACCACTACTGCGAACAAATGCAAATGGTAGTGAACTCATTCGACCAAGTAATGGGT 1101

Qy        318 IleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArg 337
Db       1102 TACGGCGCGGCGGTTCCGTACACGACATTAGCTCAAAAGGCAATGTCTAGGCATTTCCGG 1161
```

```
Qy       338 CysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGluGlu 357
Db      1162 TGTTTGAAAGACGCGGTAGCGGTTCAGCTTAAACGCAGCTGTGAGCTTCTAGGGGATAAA 1221

Qy       358 GluGlyLeuGlyGly-----------------LysIleGluGlySerArgLeuLysPhe 371
Db      1222 GAGGCGGCAGGGGCTGCATCCTCGGGGTTAACCAAAGGGGAAACGCCGCGATTGCGTTTG 1281

Qy       372 ValAspHisHisLeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGlnProAsn 391
Db      1282 CTAGAGCAGAGTTTGCGTCAGCAACGAGCGTTTCATCATATGGGTATGATGGAGCAAGAG 1341

Qy       392 AlaTrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeu 411
Db      1342 GCATGGAGACCGCAACGTGGTTTGCCTGAACGCTCCGTTAATATCCTTAGAGCTTGGCTA 1401

Qy       412 PheGluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGln 431
Db      1402 TTCGAGCATTTTCTTAATCCGTACCCAAGCGATGCTGATAAGCACCTCTTAGCACGACAG 1461

Qy       432 ThrGlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrp 451
Db      1462 ACTGGTTTATCCAGAAATCAGGTGTCAAATTGGTTCATAAATGCTAGGGTTCGCCTATGG 1521

Qy       452 LysProMetValGluGluMetTyrLeuGluGluValLys---------------------- 464
Db      1522 AAACCAATGGTGGAAGAGATGTATCAACAAGAAGCAAAAGAAAGAGAAGAAGCAGAAGAA 1581

Qy       465 ------------AsnGlnGluGlnAsnSerThrAsnThrSerGlyAspAsnLysAsnLys 480
Db      1582 GAAAATGAAAATCAACAACAACAAAGAAGACAGCAACAAACAAACAACAACGACACGAAA 1641

Qy       481 GluThrAsnIleSerAlaProAsnGluGluLysHisProIleIleThrSerSerLeuLeu 500
Db      1642 CCCAACAAC------------AATGAAAACAACTTCACTGTCATAACCGCACAA------ 1683

Qy       501 GlnAspGlyIleThrThrThrGlnAlaGluIleSerThrSerThrIleSerThrSerPro 520
Db      1684 ----------------------------------------------------ACTCCA 1689

Qy       521 ThrAlaGlyAlaSerLeuHisHisAlaHisAsnPheSerPheLeuGlySerPheAsnMet 540
Db      1690 ACGACGATGACATCGACACATCAC---------------------------------- 1713

Qy       541 AspAsnThrThrThrThrValAspHisIleGluAsnAsnAlaLysLysGlnArgAsnAsp 560
Db      1714 ----------------------------------------------------GAAAACGAC 1722

Qy       561 MetHisLysPheSerProSerSerIleLeuSerSerValAspMetGluAlaLysAlaArg 580
Db      1723 -----------------TCTTCATTCCTCTCTTCCGTC---------GCCGCCGCTTCT 1755

Qy       581 GluSerSerAsnLysGlyPheThrAsnProLeuMetAlaAlaTyrAlaMetGlyAspPhe 600
Db      1756 CACGGCGGTTCAGACGCGTTCACCGTCGCCACGTGT----------------------- 1791

Qy       601 GlyArgPheAspProHisAspGlnGlnMetThrAlaAsnPheHis------------Gly 616
Db      1792 --------------------CAGCAAGACGTCAGTGACTTCCACGTCGACGGAGATGGT 1830

Qy       617 AsnAsn----------------------------GlyValSerLeuThrLeuGlyLeu 626
Db      1831 GTGAACGTCATAAGATTCGGGACCAAACAGACTGGTGACGTGTCTCTTACGCTTGGTCTA 1890

Qy       627 ProProSerGluAsnLeuAlaMetProValSerGlnGlnAsnTyrLeuSerAsnAspLeu 646
Db      1891 CGCCACTCTGGCAATATT------CCTGATAAGAACACTTCTTTCTCCGTTAGAGACTTT 1944

Qy       647 Gly 647
Db      1945 GGA 1947

RESULT 13
```

```
ADO02224
ID   ADO02224 standard; cDNA; 1983 BP.
XX
AC   ADO02224;
XX
DT   01-JUL-2004  (first entry)
XX
DE   Thalecress transcription factor cDNA #319.
XX
KW   Thalecress; transcription factor; ss; gene; plant; transgenic;
KW   abiotic stress; cold tolerance; heat tolerance; drought; osmotic stress;
KW   phosphate limitation; potassium limitation; nitrogen limitation;
KW   hormone sensitivity; disease resistance; sugar sensing; seed germination;
KW   flowering; inflorescence architectural change;
KW   meristem cell differentiation; phyllotaxy; apical dominance;
KW   trichome development; seed development; premature senescence;
KW   delayed senescence; lethality; necrosis; plant size; leaf morphology;
KW   seed morphology; secondary metabolism; light response; shade avoidance.
XX
OS   Arabidopsis thaliana.
XX
PN   US2004045049-A1.
XX
PD   04-MAR-2004.
XX
PF   10-APR-2003; 2003US-00412699.
XX
PR   13-SEP-1999;   99US-00394519.
PR   21-JAN-2000; 2000US-00489376.
PR   17-FEB-2000; 2000US-00506720.
PR   22-MAR-2000; 2000US-00532591.
PR   22-MAR-2000; 2000US-00533029.
PR   22-MAR-2000; 2000US-00533030.
PR   22-MAR-2000; 2000US-00533392.
PR   22-MAR-2000; 2000US-00533648.
PR   06-APR-2000; 2000WO-US009448.
PR   16-NOV-2000; 2000US-00713994.
PR   27-MAR-2001; 2001US-00819142.
PR   17-APR-2001; 2001US-00837444.
PR   30-JAN-2002; 2002US-00958131.
PR   14-JUN-2002; 2002US-00171468.
PR   09-AUG-2002; 2002US-00225066.
PR   09-AUG-2002; 2002US-00225067.
PR   09-AUG-2002; 2002US-00225068.
PR   17-DEC-2002; 2002US-0434166P.
PR   25-FEB-2003; 2003US-00374780.
XX
PA   (ZHAN/) ZHANG J.
PA   (FROM/) FROMM M E.
PA   (HEAR/) HEARD J E.
PA   (RIEC/) RIECHMANN J L.
PA   (ADAM/) ADAM L J.
PA   (BROU/) BROUN P E.
PA   (PINE/) PINEDA O.
PA   (REUB/) REUBER T L.
PA   (KEDD/) KEDDIE J S.
PA   (YUGG/) YU G.
PA   (JIAN/) JIANG C.
PA   (SAMA/) SAMAHA R S.
PA   (PILG/) PILGRIM M L.
PA   (CREE/) CREELMAN R A.
PA   (DUBE/) DUBELL A N.
PA   (RATC/) RATCLIFFE O.
PA   (KUMI/) KUMIMOTO R.
PA   (SHER/) SHERMAN B K.
XX
PI   Zhang J,  Fromm ME,  Heard JE,  Riechmann JL,  Adam LJ,  Broun PE;
PI   Pineda O,  Reuber TL,  Keddie JS,  Yu G,  Jiang C,  Samaha RS;
PI   Pilgrim ML,  Creelman RA,  Dubell AN,  Ratcliffe O,  Kumimoto R;
PI   Sherman BK;
XX
DR   WPI; 2004-225755/21.
DR   P-PSDB; ADO02225.
```

XX
PT New transgenic plant, useful in developing phenotypes with altered or
PT improved characteristics or traits.
XX
PS Claim 1; SEQ ID NO 637; 213pp; English.
XX
CC The invention relates to a transgenic plant comprises a recombinant
CC polynucleotide having a polynucleotide sequence or its complementary
CC sequence comprising a sequence encoding a polypeptide, that initiates
CC transcription (i.e. a transcription factor) from Arabidopsis, Soybean,
CC Rice, Rape or Corn, comprising any of the sequences appearing as ADO01588
CC -ADO03527 or ADO03530-ADO03559. Also included are using a transgenic
CC plant to grow a progeny plant, an expression cassette (comprising a
CC constitutive, inducible or tissue-specific promoter and a recombinant
CC polynucleotide described above), a host cell comprising the expression
CC cassette, producing a modified plant having a modified trait, identifying
CC a factor that is modulated by or interacts with a polypeptide encoded by
CC the polynucleotide sequence and identifying at least one downstream
CC polynucleotide sequence that is subject to a regulatory effect of any of
CC the polypeptides encoded by the polynucleotide described above. The
CC transgenic plant is useful for producing a plant that has an altered
CC trait e.g. an enhanced tolerance to abiotic stress (increased tolerance
CC to chilling, germination in cold conditions,freezing tolerance, tolerance
CC to heat, tolerance to drought, tolerance to osmotic stress, tolerance to
CC salt, tolerance to phosphate limitation, tolerance to potassium
CC limitation, decreased sensitivity to nitrogen limitation), altered
CC hormone sensitivity, reduced sensitivity to abscisic acid, an altered
CC response to ethylene, disease resistance, altered susceptibility to
CC Botrytis, altered susceptibility to Fusarium, altered susceptibility to
CC Erysiphe, altered susceptibility to Pseudomonas syringae, altered
CC susceptibility to Sclerotinia, altered sugar sensing, improved seed
CC germination and seedling vigor, early flowering, late flowering, extended
CC period of flowering, an inflorescence architectural change, a change in
CC stem bifurcations, a lack of a shoot meristem, reduced meristem cell
CC differentiation, altered phyllotaxy, altered branching pattern, reduced
CC apical dominance, reduced trichome density, ectopic trichome development,
CC altered trichome development, altered stem morphology, increased root
CC growth, increased root hairs, altered seed development, altered cell
CC proliferation/cell differentiation, premature senescence, delayed
CC senescence, lethality, increased necrosis, an increase in seedling or
CC plant size, decreased plant size, a change in leaf morphology, increased
CC altered leaf development, increased leaf size and mass, glossy leaves,
CC leaf cell expansion, change in seed morphology, altered seed coloration,
CC increased seed size, decreased seed size, altered seed shape, change in
CC leaf biochemistry, increased leaf wax, an alteration in leaf prenyl lipid
CC content, increased leaf insoluble sugars, decreased leaf insoluble
CC sugars, increased leaf anthocyanins, an alteration of leaf fatty acid
CC content, an alteration of leaf glucosinolate content, change in seed
CC biochemistry, an increase in seed oil content, decrease in seed oil
CC content, increase in seed fatty acid content, decrease in seed fatty acid
CC content, increase in seed protein content, decrease in seed protein
CC content, alteration in seed prenyl lipid content, increase in seed
CC sterols, upregulation of genes involved in secondary metabolism, increase
CC in root anthocyanins, increase in plant anthocyanins, and alteration in
CC light response or shade avoidance. The present sequence encodes a
CC thalecress transcription factor of the invention.
XX
SQ Sequence 1983 BP; 585 A; 464 C; 434 G; 500 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.72e-58	Length:	1983
Score:	824.00	Matches:	237
Percent Similarity:	45.96%	Conservative:	76
Best Local Similarity:	34.80%	Mismatches:	178
Query Match:	23.13%	Indels:	190
DB:	12	Gaps:	25

US-10-624-201A-2 (1-688) x ADO02224 (1-1983)

```
Qy        15 AspHisGlnGlnArgHisAsnHisGlyAsnSerAsnAsnAsnAsnIleGlnThrLeuTyr 34
             :::  |||  ||||||              ||||||||||||     ||||||:::
Db       331 GAGCAACAAAACCGTCAC--------------AACAACAACAATAACTCAACTCTTCAT 375

Qy        35 LeuMetAsnProAsnAsnTyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGlnGln 54
             ::::::  ||||||::::::  ||||||:::     :::|||           |||
Db       376 ATGTTATTACCAAATCATCAT---CAAGGTTTTGCTTTCACCGACGAAAACACTATGCAG 432

Qy        55 GlnLeuLeuPheLeuAsnSerSerProAlaAlaSerAsnAlaLeuCysHisAlaAsnIle 74
Db       432 ------------------------------------------------------------ 432

Qy        75 GlnHisAlaProLeuGlnGlnGlnHisPheValGlyValProLeuProAlaValSerLeu 94
                 |||  ||||||||||||||||            |||:::  |||
Db       433 ---------CCGCAGCAACAACAACACTTTACA---------TGGCCATCTTCCTCC--- 471

Qy        95 HisAspGlnIleAsnHisHisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnSer 114
                    :::||||||                    |||:::|||
Db       472 ---------TCCGATCATCAT--------------------CAAAACCGAGATATGATC 501

Qy       115 GlnGlnValIleValProSerSerThrGlyValSerAlaThrSerCysGlyGlyIleThr 134
                 |||  |||          |||:::|||  :::          ::::::
Db       502 GGAACCGTCCACGTGGAAGGAGGAAAGGGTTTGTCTTTATCT-----------CTCTCA 549

Qy       135 ThrAspLeuAlaSerGlnLeuAlaPheGln---ArgProIle----------------- 147
                 :::  ||||||:::      |||  :::  |||  |||
Db       550 TCTTCATTAGCCGCAGCTAAAGCCGAGGAATATAGAAGCATTTATTGTGCAGCCGTTGAT 609

Qy       148 -----------------ProThrProGlnHisArgGlnGlnGlnGlnGlnGlyGly 161
                                        |||      |||      |||
Db       610 GGAACTTCTTCTTCTTCTAACGCATCCGCTCATCATCATCAATTCAATCAGTTCAAGAAT 669

Qy       162 LeuSerLeu------SerLeuSerProGlnLeuGlnGlnGlnIleSerPheAsnAsnAsn 179
                 |||  |||      |||  |||               |||:::      :::
Db       670 CTTCTTCTTGAGAATTCTTCTTCTCAACATCATCACCATCAAGTTGTTGGACATTTTGGT 729

Qy       180 IleSerSerSerSerProArgThrAsnAsnValThrIleArgGlyThrLeuAspGlySer 199
                |||||||||||||||        :::  :::|||  |||
Db       730 TCATCATCATCATCTCCCATGGCGGCTTCTTCATCCATTGGAGGGATCTAC--------- 780

Qy       200 SerSerAsnMetValLeuGlySerLysTyrLeuLysAlaAlaGlnGluLeuLeuAspGlu 219
                      :::      |||||||||  |||  |||||||||||||:::|||
Db       781 ---------ACGTTGAGGAATTCGAAATATACGAAACCGGCTCAAGAGTTGTTGGAAGAG 831

Qy       220 ValValAsnIleValGlyLysSerIleLysGlyAspAspGlnLysLysAspAsnSerMet 239
                 ::::::              |||          |||  |||  :::
Db       832 TTTTGTAGTGTT-----------------GGAAGAGGACATTTCAAGAAGAACAAACTT 873

Qy       240 AsnLysGluSerMetProLeuAlaSerAspValAsnThr------AsnSerSerGlyGly 257
                 ::::::  :::           |||:::  ||||||              ||||||
Db       874 AGTAGGAACAAC------------TCAAACCCTAATACTACCGGTGGAGGAGGAGGCGGA 921

Qy       258 GlyGluSerSerSerArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGlnArg 277
                 |||  |||||||||              :::      |||:::  |||  |||
Db       922 GGGTCCTCGTCATCGGCCGGAACAGCTAATGATAGTCCTCCTTTGTCTCCGGCTGATCGG 981

Qy       278 GlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeuGluGluValGluGlnArg 297
                 |||  |||  :::|||  |||||||||:::|||||||||||||||:::::|||
Db       982 ATTGAACATCAAAGAAGAAAAGTCAAGCTACTATCTATGCTTGAAGAGGTGGACCGACGG 1041

Qy       298 TyrArgGlnTyrHisHisGlnMetGlnIleIleValLeuSerPheGluGlnValAlaGly 317
                 |||      |||      |||||||||:::::|||   ||||||:::||||||   |||
Db      1042 TACAACCACTACTGCGAACAAATGCAAATGGTAGTGAACTCATTCGACCAAGTAATGGGT 1101

Qy       318 IleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArg 337
                 |||:::|||      ||||||  ||||||      |||:::|||:::  ||||||
Db      1102 TACGGCGCGGCGGTTCCGTACACGACATTAGCTCAAAAGGCAATGTCTAGGCATTTCCGG 1161

Qy       338 CysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGluGlu 357
                 ||||||||||||||:::|||  |||:::|||  :::  :::  ||||||::::::
Db      1162 TGTTTGAAAGACGCGGTAGCGGTTCAGCTTAAACGCAGCTGTGAGCTTCTAGGGGATAAA 1221

Qy       358 GluGlyLeuGlyGly-----------------LysIleGluGlySerArgLeuLysPhe 371
                 |||      |||                |||  |||  ||||||:::
Db      1222 GAGGCGGCAGGGGCTGCATCCTCGGGGTTAACCAAAGGGGAAACGCCGCGATTGCGTTTG 1281

Qy       372 ValAspHisHisLeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGlnProAsn 391
```

```
            ::::::      ||||||||||||||||||        :::|||||||||:::
Db       1282 CTAGAGCAGAGTTTGCGTCAGCAACGAGCGTTTCATCATATGGGTATGATGGAGCAAGAG 1341

Qy        392 AlaTrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeu 411
               ||||||||||||||||||||||||||||||:::|||::::::|||||||||||||||
Db       1342 GCATGGAGACCGCAACGTGGTTTGCCTGAACGCTCCGTTAATATCCTTAGAGCTTGGCTA 1401

Qy        412 PheGluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGln 431
               ||||||||||||||:::||||||||||   |||:::||||||   :::||||||:::|||
Db       1402 TTCGAGCATTTTCTTAATCCGTACCCAAGCGATGCTGATAAGCACCTCTTAGCACGACAG 1461

Qy        432 ThrGlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrp 451
               |||||||||:::|||:::|||||||||||||||||||||||||||||||||||||||||
Db       1462 ACTGGTTTATCCAGAAATCAGGTGTCAAATTGGTTCATAAATGCTAGGGTTCGCCTATGG 1521

Qy        452 LysProMetValGluGluMetTyrLeuGluGluValLys-------------------- 464
               ||||||||||||||||||||||   :::|||   |||
Db       1522 AAACCAATGGTGGAAGAGATGTATCAACAAGAAGCAAAAGAAAGAGAAGAAGCAGAAGAA 1581

Qy        465 ------------AsnGlnGluGlnAsnSerThrAsnThrSerGlyAspAsnLysAsnLys 480
                         |||:::|||               :::   :::|||      |||
Db       1582 GAAAATGAAAATCAACAACAACAAAGAAGACAGCAACAAACAAACAACAACGACACGAAA 1641

Qy        481 GluThrAsnIleSerAlaProAsnGluGluLysHisProIleIleThrSerSerLeuLeu 500
               |||             ||||||            :::||||||:::
Db       1642 CCCAACAAC------------AATGAAAACAACTTCACTGTCATAACCGCACAA------ 1683

Qy        501 GlnAspGlyIleThrThrThrGlnAlaGluIleSerThrSerThrIleSerThrSerPro 520
                                                                  :::|||
Db       1684 ----------------------------------------------------ACTCCA 1689

Qy        521 ThrAlaGlyAlaSerLeuHisHisAlaHisAsnPheSerPheLeuGlySerPheAsnMet 540
               |||         |||   ||||||
Db       1690 ACGACGATGACATCGACACATCAC---------------------------------- 1713

Qy        541 AspAsnThrThrThrThrValAspHisIleGluAsnAsnAlaLysLysGlnArgAsnAsp 560
                                                                  ||||||
Db       1714 ------------------------------------------------GAAAACGAC 1722

Qy        561 MetHisLysPheSerProSerSerIleLeuSerSerValAspMetGluAlaLysAlaArg 580
                          ||||||   ||||||||||||         |||   |||
Db       1723 -----------------TCTTCATTCCTCTCTTCCGTC---------GCCGCCGCTTCT 1755

Qy        581 GluSerSerAsnLysGlyPheThrAsnProLeuMetAlaAlaTyrAlaMetGlyAspPhe 600
                     :::      ||||||
Db       1756 CACGGCGGTTCAGACGCGTTCACCGTCGCCACGTGT----------------------- 1791

Qy        601 GlyArgPheAspProHisAspGlnGlnMetThrAlaAsnPheHis-----------Gly 616
                              ||||||      :::::||||||  .          |||
Db       1792 --------------------CAGCAAGACGTCAGTGACTTCCACGTCGACGGAGATGGT 1830

Qy        617 AsnAsn----------------------------GlyValSerLeuThrLeuGlyLeu 626
                  |||                          ||||||||||||||||||||
Db       1831 GTGAACGTCATAAGATTCGGGACCAAACAGACTGGTGACGTGTCTCTTACGCTTGGTCTA 1890

Qy        627 ProProSerGluAsnLeuAlaMetProValSerGlnGlnAsnTyrLeuSerAsnAspLeu 646
                     |||   |||:::      |||            ::::::         |||
Db       1891 CGCCACTCTGGCAATATT------CCTGATAAGAACACTTCTTTCTCCGTTAGAGACTTT 1944

Qy        647 Gly 647
               |||
Db       1945 GGA 1947

RESULT 14
AAC39138
ID   AAC39138 standard; DNA; 2025 BP.
XX
AC   AAC39138;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 23512.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway; metabolic pathway;
KW   promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-00301439.
XX
PR   25-FEB-1999;   99US-0121825P.
PR   05-MAR-1999;   99US-0123180P.
PR   09-MAR-1999;   99US-0123548P.
PR   23-MAR-1999;   99US-0125788P.
PR   25-MAR-1999;   99US-0126264P.
PR   29-MAR-1999;   99US-0126785P.
PR   01-APR-1999;   99US-0127462P.
PR   06-APR-1999;   99US-0128234P.
PR   08-APR-1999;   99US-0128714P.
PR   16-APR-1999;   99US-0129845P.
PR   19-APR-1999;   99US-0130077P.
PR   21-APR-1999;   99US-0130449P.
PR   23-APR-1999;   99US-0130510P.
PR   23-APR-1999;   99US-0130891P.
PR   28-APR-1999;   99US-0131449P.
PR   30-APR-1999;   99US-0132048P.
PR   30-APR-1999;   99US-0132407P.
PR   04-MAY-1999;   99US-0132484P.
PR   05-MAY-1999;   99US-0132485P.
PR   06-MAY-1999;   99US-0132486P.
PR   06-MAY-1999;   99US-0132487P.
PR   07-MAY-1999;   99US-0132863P.
PR   11-MAY-1999;   99US-0134256P.
PR   14-MAY-1999;   99US-0134218P.
PR   14-MAY-1999;   99US-0134219P.
PR   14-MAY-1999;   99US-0134221P.
PR   14-MAY-1999;   99US-0134370P.
PR   18-MAY-1999;   99US-0134768P.
PR   19-MAY-1999;   99US-0134941P.
PR   20-MAY-1999;   99US-0135124P.
PR   21-MAY-1999;   99US-0135353P.
PR   24-MAY-1999;   99US-0135629P.
PR   25-MAY-1999;   99US-0136021P.
PR   27-MAY-1999;   99US-0136392P.
PR   28-MAY-1999;   99US-0136782P.
PR   01-JUN-1999;   99US-0137222P.
PR   03-JUN-1999;   99US-0137528P.
PR   04-JUN-1999;   99US-0137502P.
PR   07-JUN-1999;   99US-0137724P.
PR   08-JUN-1999;   99US-0138094P.
PR   10-JUN-1999;   99US-0138540P.
PR   10-JUN-1999;   99US-0138847P.
PR   14-JUN-1999;   99US-0139119P.
PR   16-JUN-1999;   99US-0139452P.
PR   16-JUN-1999;   99US-0139453P.
PR   17-JUN-1999;   99US-0139492P.
PR   18-JUN-1999;   99US-0139454P.
PR   18-JUN-1999;   99US-0139455P.
PR   18-JUN-1999;   99US-0139456P.
PR   18-JUN-1999;   99US-0139457P.
PR   18-JUN-1999;   99US-0139458P.
PR   18-JUN-1999;   99US-0139459P.
PR   18-JUN-1999;   99US-0139460P.
PR   18-JUN-1999;   99US-0139461P.
PR   18-JUN-1999;   99US-0139462P.
PR   18-JUN-1999;   99US-0139463P.
PR   18-JUN-1999;   99US-0139750P.
PR   18-JUN-1999;   99US-0139763P.
PR   21-JUN-1999;   99US-0139817P.
PR   22-JUN-1999;   99US-0139899P.
PR   23-JUN-1999;   99US-0140353P.
```

PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:

Pred. No.:	3.8e-58	Length:	2025
Score:	820.00	Matches:	222
Percent Similarity:	49.84%	Conservative:	92
Best Local Similarity:	35.24%	Mismatches:	143
Query Match:	23.01%	Indels:	173
DB:	3	Gaps:	23

US-10-624-201A-2 (1-688) x AAC39138 (1-2025)

```
Qy         16 HisGlnGlnArgHisAsn----------------HisGlyAsnSerAsnAsnAsnAsn 29
              |||      :::||||||                      ||||||:::|||
Db         95 CACACTCACAAACACAATCTCTAGGATCTCATTTCCCGACCAAACTCATCGAACTAAGAA 154

Qy         30 IleGlnThrLeuTyrLeuMetAsnProAsnAsnTyr-MetGlnGlyTyrThrThrSerAs 49
              :::      :::   |||:::         |||    |||      |||   ||||||::
Db        155 CTCACTGTAATAAAACTTTTAAAAGGTGTAAACACCAATGGCAGTTTATTACACAAGTAA 214

Qy         49 pThrGln---GlnGlnGlnGlnLeuLeuPheLeu---------------AsnSerSerPr 63
              :          ||||||:::   :::||||||               |||
Db        215 TGTCGGTTGTTACCAGCAAGAACCAATCTTTCTCAACCATCAACAACAAAACCAACAAGC 274

Qy         63 oAlaAlaSerAsnAlaLeuCysHisAlaAsnIleGlnHisAlaProLeuGlnGlnGlnHi 83
               ::::::|||:::|||             :::                    :::
Db        275 TTCTTCTTCATCCGCCGCCGCTTCTTTCACCGTCACCGGCGGCGATACTGTTCGAAACGA 334
```

```
Qy        83 sPheValGlyValProLeuProAlaValSerLeuHisAspGlnIleAsnHisHisGlyLe 103
Db       335 GATGGTTTTTATCCCA---------------------------------------------- 350

Qy       103 uLeuGlnArgMetTrpAsnAsnGlnAspGlnSerGlnGlnValIleValProSerSerTh 123
Db       351 ------------------------------------------------CCAACCACCAC 361

Qy       123 rGlyValSerAlaThrSerCysGlyGlyIleThrThrAspLeuAlaSerGlnLeuAlaPh 143
Db       362 AGGAGACGTTGTAACAGGAAACGGTACCGTTTCA---------AGCAGCGATCTAAGCTT 412

Qy       143 eGlnArgProIleProThrProGlnHisArgGlnGlnGlnGlnGlnGlnGlyGlyLeuSe 163
Db       413 T-----------------------CACGATGGTCAA--------------GGACTGTC 433

Qy       163 rLeuSerLeuSerProGlnLeu-----------------------GlnGlnGlnIleSe 175
Db       434 TTTGAGCCTTGGTACTCAGATCTCTGTTGCTCCGTTTCACTTTCATCAATACCAATTGGG 493

Qy       175 rPheAsnAsn--------------AsnIleSerSerSerSerPro-------------- 185
Db       494 GTTTACTAGTCAGAATCCTTCAATTTCAGTCAAGGAAACGTCACCGTTTCATGTGGATGA 553

Qy       186 ----------ArgThrAsnAsnValThrIleArgGlyThrLeuAspGlySerSerSer-- 201
Db       554 GATGAGTGTGAAGAGCAAAGAGATGATATTGTTGGGTCAATCTGATCCTTCCTCTGGTTA 613

Qy       201 ------------------------------------------------------------ 201
Db       614 TGCTGGTAATGGTGGGAATGGCTTCTACAACAATTATCGGTATAATGAGACATCAGGAGG 673

Qy       202 -------AsnMetValLeuGlySerLysTyrLeuLysAlaAlaGlnGluLeuLeuAspGl 219
Db       674 GTTTATGAGCAGCGTTCTGCGTTCTCGGTATCTTAAACCTGCTCAGAATTTGCTTGATGA 733

Qy       219 uValValAsnIleValGlyLysSerIleLysGlyAspAspGlnLysLysAspAsnSerMe 239
Db       734 AGTGGTTAGTGTC-------------------------AAGAAAGAACTAAACCAAAT 766

Qy       239 tAsnLysGluSerMetProLeuAlaSerAspValAsnThrAsnSer------------Se 255
Db       767 GAGTAAGAAGAAGATGAAAGTT---AATGACTTTAACAATGGTTCTAAGGAGATAGAAGG 823

Qy       255 rGlyGlyGlyGluSerSerSerArgGlnLysAsnGluValAlaValGluLeuThrThrAl 275
Db       824 AGGAGGCAGTGGCGAGCTATCGAATGATTCGAATGGGAAATCGATTGAGTTATCTACTAT 883

Qy       275 aGlnArgGlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeuGluGluValGl 295
Db       884 TGAACGTGATGAGCTTCATAACAAGAAGAACAAGCTTTTAACAATGGTTGATGAGGTAGA 943

Qy       295 uGlnArgTyrArgGlnTyrHisHisGlnMetGlnIleIleValLeuSerPheGluGlnVa 315
Db       944 TAAAAGATATAACCAATATTACCATCAAATGGAAGCATTAGCTTCATCATTTGAGATAGT 1003

Qy       315 lAlaGlyIleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGl 335
Db      1004 AGCAGGACTTGGATCAGCTAAGCCTTACACATCAGTTGCTCTCAACAGAATCTCTCGCCA 1063

Qy       335 nPheArgCysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGl 355
Db      1064 TTTTCGTGCTCTTCGCGACGCAATAAAGGAACAAATTCAGATTGTTAGAGAGAAACTTGG 1123

Qy       355 yGlu------------------GluGluGlyLeuGlyGlyLysIleGluGlySerArgLe 369
Db      1124 GGAGAAAGGAGGAGAGTCGTTGGATGAGCAACAAGGAGAGAGGATA------CCAAGACT 1177

Qy       369 uLysPheValAspHisHisLeuArgGlnGlnArgAlaLeu---GlnGlnIleGlyMetMe 388
Db      1178 CAGGTATTTAGATCAACGGTTGAGACAGCAAAGAGCTTTGCATCAACAGCTTGGAATGGT 1237

Qy       388 tGlnProAsnAlaTrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeuAr 408
Db      1238 TCGACCC---GCTTGGAGACCTCAAAGAGGTCTTCCTGAAAACTCTGTTTCTGTTCTCCG 1294

Qy       408 gAlaTrpLeuPheGluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLe 428
Db      1295 TGCTTGGCTCTTCGAACATTTCCTTCATCCATATCCGAAAGAATCTGAGAAAATCATGCT 1354

Qy       428 uAlaLysGlnThrGlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgVa 448
Db      1355 TGCGAAACAGACAGGACTGTCGAAGAACCAGGTTGCTAATTGGTTCATAAACGCGAGAGT 1414

Qy       448 lArgLeuTrpLysProMetValGluGluMetTyrLeuGluGluVal-------------- 463
Db      1415 TCGTCTATGGAAACCGATGATCGAAGAGATGTATAAAGAAGAGTTTGGTGATGAATCTGA 1474

Qy       464 ----------------LysAsnGlnGluGlnAsnSerThrAsnThrSerGlyAspAsn-- 477
Db      1475 GTTACTAATCTCCAAATCTTCCCAAGAACCCAACAGCACAAACCAAGAAGACTCCTCATC 1534

Qy       478 ----------------------LysAsnLysGluThrAsnIleSerAlaProAsnGluGl 490
Db      1535 GCAGCAGCAGCAGCAGCAAGAGAACAACAACAACAGCAACCTCGCTTATTCATCTGCAGA 1594

Qy       490 uLysHisProIleIleThrSerSer------------------LeuLeuGlnAspGlyIl 504
Db      1595 CACAACAAACATTGTCTTCTCATCAGAAACCAAACCAGATCGTGTTCTAGGCAATGATAA 1654

Qy       504 eThrThrThrGlnAlaGluIleSerThrSerThr-------------------------- 515
Db      1655 CGACCCACAGCAACAACAGATAAACCGGTCATCGGATTACGACACTCTGATGAACTATCA 1714

Qy       516 ----------------------------IleSerThrSer----------ProThrAlaGl 523
Db      1715 CGGGTTTGGTGTTGATGATTACCGTTACATTAGCGGAAGCAACCAGCAAGAAAGTAGATT 1774

Qy       523 yAlaSerLeuHisHisAlaHisAsnPhe 532
Db      1775 CTCCAATTCCCATCACTTACACGACTTT 1802
```

RESULT 15
ADO62707
ID ADO62707 standard; DNA; 2458 BP.
XX
AC ADO62707;
XX
DT 15-JUL-2004 (first entry)
XX
DE Transcription factor G2550 orthologous sequence, SEQ ID 1174.
XX
KW Plant; transcription factor; transgenic plant; abiotic stress tolerance;
KW osmotic stress tolerance; cold tolerance; heat tolerance;
KW low nitrogen tolerance; low phosphate tolerance; fungal disease;
KW glyphosate resistance; flowering; fertility; seed development; ds.
XX
OS Glycine max.
XX
PN WO2004031349-A2.
XX
PD 15-APR-2004.
XX
PF 18-SEP-2003; 2003WO-US030292.
XX
PR 18-SEP-2002; 2002US-0411837P.
PR 17-DEC-2002; 2002US-0434166P.
PR 24-APR-2003; 2003US-0465809P.
XX
PA (MEND-) MENDEL BIOTECHNOLOGY INC.
XX
PI Jiang C, Heard JE, Ratcliffe O, Creelman RA, Adam LJ, Reuber TL;
PI Riechmann JL, Haake V, Dubell AN, Keddie JS, Sherman BK;
XX
DR WPI; 2004-330163/30.
XX

```
PT   New recombinant polynucleotide encoding transcription factor
PT   polypeptides, useful for producing transgenic plants with advantageous
PT   properties compared to a reference plant.
XX
PS   Claim 1; SEQ ID NO 1174; 510pp; English.
XX
CC   The present invention relates to novel plant transcription factor
CC   proteins (I) and nucleotide sequences (II) (ADO61534-ADO63778). The
CC   sequences can be used to produce transgenic plants, which overexpress
CC   (II), where the transgenic plant has an altered trait as compared to a
CC   non-transgenic plant or wild-type plant. The transgenic plant comprises
CC   an altered trait selected from increased tolerance to abiotic stress,
CC   increased tolerance to osmotic stress, increased tolerance to cold,
CC   increased germination in cold, increased tolerance to heat, increased
CC   germination in heat, increased tolerance to freezing conditions,
CC   increased tolerance to low nitrogen conditions, increased tolerance to
CC   low phosphate conditions, increased tolerance to disease, including
CC   fungal disease and particularly Erysiphe, Fusarium and Botrytis,
CC   increased tolerance to multiple fungal pathogens, increased resistance to
CC   glyphosate, increased sensitivity to ABA, reduced sensitivity to ABA,
CC   increased sensitivity to ACC, altered sugar sensing, increased tolerance
CC   to sugars, altered carbon/nitrogen sensing, early flowering, late
CC   flowering, altered flower structure, loss of flower determinacy, reduced
CC   fertility, altered shoot meristem development, altered branching pattern,
CC   altered stem morphology, altered vascular tissue structure, reduced
CC   apical dominance, altered trichome density, altered trichome development,
CC   altered trichome structure, altered root development, altered shade
CC   avoidance, altered seed development, altered seed ripening, altered seed
CC   germination, slow growth, fast growth, altered cell differentiation,
CC   altered cell proliferation, altered cell expansion, altered phase change,
CC   altered senescence, abnormal embryo development, altered programmed cell
CC   death, lethality when overexpressed, altered necrosis patterns, increased
CC   plant size, increased biomass, large seedlings, dwarfed plants, dark
CC   green leaves, change in leaf shape, increased leaf size and mass, light
CC   green or gray leaves, glossy leaves, altered abaxial/adaxial polarity,
CC   altered seed coloration, altered seed size, altered seed shape, large
CC   seed, increased leaf wax, increased leaf fatty acids, altered seed oil
CC   content, altered seed protein content, altered seedprenyl content,
CC   altered leaf prenyl lipid content, increased anthocyanin levels, and
CC   decreased anthocyanin levels. Note: The sequence data for this patent did
CC   not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 2458 BP; 751 A; 488 C; 545 G; 674 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                1.06e-57      Length:          2458
Score:                    816.00        Matches:         218
Percent Similarity:       49.92%        Conservative:    83
Best Local Similarity:    36.15%        Mismatches:      149
Query Match:              22.90%        Indels:          153
DB:                       12            Gaps:            21

US-10-624-201A-2 (1-688) x ADO62707 (1-2458)

Qy         69 LeuCysHisAlaAsnIleGlnHisAlaProLeuGln---GlnGlnHisPheValGlyVal 87
              |||      :::|||         :::|||   |||   |||   |||      |||:::
Db        513 CTCATGCCTTCTAATGCATCCGTCTCTCCATTTCAATACCAATATCACGACACGGGTCTC 572

Qy         88 --------------------------ProLeuProAlaValSerLeuHisAspGln--- 97
                                        |||      |||:::||||||   |||
Db        573 TTACCACTAATGAATGATTTTCCAAATCCAAAGGGAGCCATGTCTCTCAAAGATGATGAG 632

Qy         98 ---IleAsnHisHisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnSerGlnGln 116
                 |||   ||||||      :::
Db        633 GCCAATAACCTACATGGGGAGTTTAGA------------------------------- 659

Qy        117 ValIleValProSerSerThrGlyValSerAlaThrSerCysGlyGlyIleThrThrAsp 136
                  |||:::     :::::::::   |||   ||||||          |||
Db        660 -----------AGTGCTGAATGCATGGCATCTGTGTCTTCTGGAGGCTTT---CTTGAC 704

Qy        137 LeuAlaSerGlnLeuAlaPheGlnArgProIleProThrProGlnHisArgGlnGlnGln 156
```

```
             :::      :::      |||      |||              |||
Db        705 ATGGTAAAAAAGGAGTGTTTTTACAACCCC--------------CATGACCCC------ 743

Qy        157 GlnGlnGlnGlyGlyLeuSerLeuSerLeuSerProGlnLeuGlnGlnGlnIleSerPhe 176
                                  |||:::  |||
Db        744 -----------------TCAATGTGTCTT---------------------------- 755

Qy        177 AsnAsnAsnIleSerSerSerSerProArgThrAsnAsnValThrIleArgGlyThrLeu 196
                    :::   |||      ||||||
Db        756 ---AAAGAAGTCCCCTCTGATCTACCACGT---------------------------- 782

Qy        197 AspGlySerSerSerAsnMetValLeuGlySerLysTyrLeuLysAlaAlaGlnGluLeu 216
                       ||||||   :::|||   |||:::||||||||||||||||:::|||
Db        783 ---------TACTCAAACAGTATTTTGAACTCTCAATACCTGAAGGCGGCACAAGATTTG 833

Qy        217 LeuAspGluValValAsnIleValGlyLysSerIleLysGlyAspAspGlnLys------ 234
                 |||||||||:::                        |||      :::|||
Db        834 CTTGATGAAATA-------------------------GGTAGTGGCCGAAAGGCTTTG 866

Qy        235 LysAspAsnSerMetAsnLysGluSerMetProLeuAlaSerAspValAsnThrAsnSer 254
                 |||  :::   |||   |||:::                    |||   :::   :::|||
Db        867 AAGCAATCTGGAATGGAAAAACAAGAG---AACACTGGATTAGATGGCTCTAAAGATTCT 923

Qy        255 SerGlyGlyGlyGluSerSerSerArgGlnLysAsn---------------------- 266
                  |||         |||   |||   |||   :::
Db        924 GATGGAAAATCTACCAGTCAATCTATGCAAATGTCTTCAGGCCCTAATGGTTCCACTGCT 983

Qy        267 GluValAlaValGluLeuThrThrAlaGlnArgGlnGluLeuGlnMetLysLysAlaLys 286
                  :::   ||||||:::::|||:::||||||   |||      ||||||   |||
Db        984 AATGCTTCTAGTGAACTATCATCCGCAGAGCGGCAGAACTTGTTGGACAAGAAAACAAAG 1043

Qy        287 LeuLeuAlaMetLeuGluGluValGluGlnArgTyrArgGlnTyrHisHisGlnMetGln 306
                 ||||||:::||||||:::||||||:::::::||||||||||||||   ||||||||||||
Db       1044 CTTTTGTCCATGTTGGATGAGGTAGATAAAAGATACAGACAGTACTGCCATCAGATGCAG 1103

Qy        307 IleIleValLeuSerPheGluGlnValAlaGlyIleGlySerAlaLysSerTyrThrGln 326
                 |||:::|||   ||||||:::   |||||||||   |||:::|||:::   ||||||
Db       1104 ATTGTGGTGTCATCTTTGACATGGTTGCTGGCTGTGGAGCAGCAGAACCATATACCACA 1163

Qy        327 LeuAlaLeuHisAlaIleSerLysGlnPheArgCysLeuLysAspAlaIleAlaGluGln 346
                 |||||||||      ||||||:::   ||||||||||||:::|||||||||:::   |||
Db       1164 CTTGCCTTAAGAACAATTTCTCGCCACTTTCGGTGTTTGCGTGATGCCATCAGTGGCCAA 1223

Qy        347 ValLysAlaThrSerLysSerLeuGlyGluGluGluGlyLeuGlyGlyLysIleGluGly 366
                 ::::::   |||   :::|||||||||   |||:::|||:::
Db       1224 ATTCAGGTGACCCAAAGAAGCCTTGGG---GAGCAAGGAATA----------------- 1262

Qy        367 SerArgLeuLysPheValAspHisHisLeuArgGlnGlnArgAlaLeuGlnGlnIleGly 386
                 ||||||::::::|||||      ||||||||||||:::||||||||||||:::|||
Db       1263 CCCCGTCTCCGCTATGTGGATCAGCAACTTAGACAACAAAAGGCACTTCAGCAACTTGGT 1322

Qy        387 MetMetGlnProAsnAlaTrpArgProGlnArgGlyLeuProGluArgAlaValSerVal 406
                 :::|||:::      |||||||||||||||||||||||||||||||   :::||||||:::
Db       1323 GTAATGAGA---CAAGCTTGGAGGCCTCAGAGGGGTCTTCCTGAAAGCTCTGTTTCAATA 1379

Qy        407 LeuArgAlaTrpLeuPheGluHisPheLeuHisProTyrProLysAspSerAspLysIle 426
                 ||||||||||||||||||||||||||||||||||||||||||||||||:::||||||
Db       1380 CTCCGTGCTTGGCTCTTTGAACATTTTCTTCATCCTTACCCCAAGGATTCTGAGAAAATT 1439

Qy        427 MetLeuAlaLysGlnThrGlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAla 446
                 |||||||||:::|||||||||||||||||:::||||||:::|||||||||||||||||
Db       1440 ATGCTGGCTAGGCAAACAGGCTTGACCAGAAACCAGGTGGCCAATTGGTTCATCAATGCA 1499

Qy        447 ArgValArgLeuTrpLysProMetValGluGluMetTyrLeuGluGluValLysAsnGln 466
                 ||||||||||||||||||||||||||||||||||||||||   ||||||      :::
Db       1500 AGGGTGCGTCTCTGGAAGCCAATGGTTGAGGAAATGTACAAAGAGGAGTTTGGAGACTCC 1559

Qy        467 GluGlnAsn------------------------------------------------- 469
                 |||   |||
Db       1560 GAGATGAACTGCAACCTTTCATCTGAGAACAACACGGTCAAATGTAAAAAAGATGATCTT 1619

Qy        470 ---------SerThrAsnThrSerGlyAspAsnLysAsnLysGluThrAsnIleSerAla 486
                      |||   |||   :::   :::||||||   :::||||||      :::
```

```
Db       1620 GTTCAAGAGTCTGATAACATTAACAACAATAATAAATGGGAAGAGACACAGGACAATTTA 1679

Qy        487 ProAsnGluGluLysHisProIleIleThrSerSerLeuLeuGlnAspGlyIleThrThr 506

Db       1680 GTAACTGTTGACAGTGTCCAAGTCCAACCTCAAGCAGAACTGGATAGAGGGGTCAATGTA 1739

Qy        507 ThrGlnAlaGluIleSerThrSerThr-------------------------------- 515

Db       1740 GAAAATGTGGTGATGGATTCTGGAACTGGGAAATTGCAGGGTGATCAAAACCAACAAAGG 1799

Qy        516 --------------------------------IleSerThrSerProThr--------- 521

Db       1800 CTTGGCATGAACAACAACAACAACTTTTATTCTATCTCAACCAACCAGAATGATGGAGGC 1859

Qy        522 ---AlaGlyAlaSerLeuHisHisAlaHisAsnPheSerPheLeuGlySerPheAsnMet 540

Db       1860 CTAATGGGTTGTTCTACTCATGCCACATATGATTTATCAGAGTTGGGTAACTTTACTGTG 1919

Qy        541 AspAsnThrThrThrThrValAspHisIleGluAsn----------------------- 552

Db       1920 GGTAGCCACGTTTCCCTTGCATTGGAGTTGAGGAACTGTGAAAGTGATGGGTTTGCCATG 1979

Qy        553 ---------AsnAlaLysLysGlnArgAsnAspMetHisLysPheSerProSerSerIle 569

Db       1980 TCTGATGATGCCATACGCAAACGACGTAATCAAGCATTGGCCTCTTCCCCGGAGACTGAT 2039

Qy        570 LeuSerSerValAspMetGlu-AlaLysAlaArgGluSerSerAsnLysGlyPheThrAs 589

Db       2040 TTG------CTAGATTATCATTTCACGGATTCAGGAAAGCAACAACACAGGTTTGGCAAT 2093

Qy        589 nProLeu 591

Db       2094 CCCCATT 2100
```

Search completed: July 29, 2005, 15:05:54
Job time : 1040 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 29, 2005, 14:20:37 ; Search time 307 Seconds
(without alignments)
3666.965 Million cell updates/sec

Title: US-10-624-201A-2
Perfect score: 3563
Sequence: 1 MYYQGTSDNTNIQADHQQRH.....GNKRFPTQLLPDFVTGNLGT 688

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10624201/runat_27072005_122958_10931/app_query.fasta_1.839
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10624201_@CGN_1_1_177_@runat_27072005_122958_10931 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	586.5	16.5	1239	4	US-09-640-211A-65	Sequence 65, Appl
2	493	13.8	529	4	US-09-640-211A-77	Sequence 77, Appl
3	445	12.5	381	4	US-09-640-211A-102	Sequence 102, App
4	445	12.5	381	4	US-09-640-211A-1472	Sequence 1472, Ap
5	420	11.8	316	4	US-09-640-211A-1854	Sequence 1854, Ap
6	420	11.8	462	4	US-09-640-211A-359	Sequence 359, App
7	419	11.8	343	4	US-09-640-211A-1578	Sequence 1578, Ap
8	411	11.5	323	4	US-09-640-211A-418	Sequence 418, App
9	360	10.1	698	4	US-09-640-211A-37	Sequence 37, Appl
10	328.5	9.2	690	4	US-09-640-211A-400	Sequence 400, App
11	319.5	9.0	380	4	US-09-640-211A-111	Sequence 111, App
12	319.5	9.0	380	4	US-09-640-211A-1246	Sequence 1246, Ap
13	297	8.3	366	4	US-09-640-211A-1241	Sequence 1241, Ap
14	260	7.3	171	4	US-09-640-211A-1845	Sequence 1845, Ap
15	256	7.2	260	4	US-09-640-211A-351	Sequence 351, App
16	256	7.2	260	4	US-09-640-211A-1833	Sequence 1833, Ap
17	251	7.0	474	4	US-09-640-211A-57	Sequence 57, Appl
c 18	251	7.0	474	4	US-09-640-211A-1399	Sequence 1399, Ap
19	236.5	6.6	486	4	US-09-640-211A-50	Sequence 50, Appl
20	233	6.5	158	4	US-09-640-211A-370	Sequence 370, App
21	226.5	6.4	407	4	US-09-640-211A-383	Sequence 383, App
22	213	6.0	157	4	US-09-640-211A-394	Sequence 394, App
23	209.5	5.9	2905	4	US-09-919-039-381	Sequence 381, App
24	208.5	5.9	2905	4	US-09-976-594-1031	Sequence 1031, Ap
25	208	5.8	533	4	US-09-640-211A-185	Sequence 185, App
26	201	5.6	2470	4	US-09-949-016-830	Sequence 830, App
27	198.5	5.6	3190	4	US-09-949-016-5219	Sequence 5219, Ap
28	197.5	5.5	3439	4	US-09-949-016-648	Sequence 648, App
29	197	5.5	1770	4	US-09-976-594-195	Sequence 195, App
30	188	5.3	2200	4	US-09-949-016-3863	Sequence 3863, Ap
31	174	4.9	2049	4	US-09-624-945-2	Sequence 2, Appli
32	172	4.8	273	4	US-09-313-294A-2156	Sequence 2156, Ap
33	171	4.8	1427	4	US-09-949-016-1856	Sequence 1856, Ap
34	168.5	4.7	2952	1	US-08-306-691B-54	Sequence 54, Appl
35	166.5	4.7	6854	1	US-08-468-036-4	Sequence 4, Appli
36	166.5	4.7	6854	2	US-08-376-843-4	Sequence 4, Appli
37	166	4.7	1233	4	US-09-270-767-11081	Sequence 11081, A
38	165.5	4.6	1086	4	US-09-248-796A-5149	Sequence 5149, Ap
39	163	4.6	1530	4	US-09-949-016-4288	Sequence 4288, Ap
40	163	4.6	1562	4	US-09-949-016-460	Sequence 460, App
41	162.5	4.6	1579	4	US-09-620-312D-927	Sequence 927, App
42	161	4.5	2965	4	US-09-665-479A-13	Sequence 13, Appl
43	160	4.5	440	4	US-09-640-211A-94	Sequence 94, Appl
44	155.5	4.4	282	4	US-09-313-294A-2142	Sequence 2142, Ap
45	154.5	4.3	2537	4	US-09-949-016-4397	Sequence 4397, Ap

ALIGNMENTS

RESULT 1
US-09-640-211A-65
; Sequence 65, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; TITLE OF INVENTION: Modification of Gene Transcription
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 65
; LENGTH: 1239
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-65

Alignment Scores:

Pred. No.:	1.7e-56	Length:	1239
Score:	586.50	Matches:	151
Percent Similarity:	50.39%	Conservative:	42
Best Local Similarity:	39.43%	Mismatches:	97
Query Match:	16.46%	Indels:	93
DB:	4	Gaps:	12

US-10-624-201A-2 (1-688) x US-09-640-211A-65 (1-1239)

```
Qy         20 HisAsnHisGlyAsnSerAsnAsnAsnAsnIleGlnThrLeuTyrLeuMetAsnProAsn 39
              |||         :::       :::       :::|||||||||   ||||||||||||
Db        240 CACCAGAGCTCATCTGACCCAGATGGAGCCCTACAAACCCTCGTCCTCATGAACCCCGCC 299
```

```
Qy        40 AsnTyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGlnGlnLeuLeuPheLeu 59
Db       300 AGCTACGTCCACTACTCCGATGCCCCGCCTCCGCACCAGCAACCCTCGGCGATCTTCCTC 359

Qy        60 AsnSerSerProAlaAlaSerAsnAlaLeuCysHisAlaAsnIleGlnHisAlaProLeu 79
Db       360 AACTCCTCCACCGCCGGGCCGCCCGCA----------------------------TCC 389

Qy        80 GlnGlnGlnHisPheValGlyValProLeuProAlaValSerLeuHisAspGln------ 97
Db       390 CAGACCCAGCAATTTGTAGGCATCCCCCTACCCGGCAGCGCCGCCGACTCGCAGCCCTCG 449

Qy        98 ---------IleAsnHis--------------HisGlyLeuLeuGlnArg--------- 106
Db       450 TCCATGCACGTCAACCACGATCTCTCCTCCATGCATGGCTTCATGCCTCGCGTCCAGTAT 509

Qy       107 ---MetTrpAsnAsnGlnAspGlnSerGlnGlnValIleValProSerSerThrGlyVal 125
Db       510 AACCTCTGGAGCTCCCTCGAC--------------------CCGTCCACGGCGGCGCGT 548

Qy       126 SerAlaThrSerCysGlyGlyIleThrThrAspLeuAlaSerGlnLeuAlaPheGlnArg 145
Db       549 GAGGCCTCC---------------------------------------------------- 557

Qy       146 ProIleProThrProGlnHisArgGlnGlnGlnGlnGlnGlnGlyGlyLeuSerLeuSer 165
Db       558 --------------------CGCACCCACCAGCAGCAG------GGGCTCTCCCTTAGC 590

Qy       166 LeuSerProGlnLeuGlnGlnGlnIle--------------SerPhe------------ 176
Db       591 CTCTCCCCGCAGCAGCCCCCACCGACCCCAGCTGGGTACCGTTCCTTTGTCCGGGCCGAG 650

Qy       177 ------------------------------------------------------Asn 177
Db       651 CGCAGTGGCGATGGAGCAGCGGGTTCTCAGCACCCTCCAGCAATTTCAGGCGGTGAGGAC 710

Qy       178 AsnAsnIleSerSerSerSerProArgThrAsnAsnValThrIleArgGlyThrLeuAsp 197
Db       711 GTGCGGATCTCGGGCGGGTCCCCATCGTCGGCCTCAGGCGTAACCAATGGGGCGGCAGTT 770

Qy       198 GlySerSer---SerAsnMetValLeuGlySerLysTyrLeuLysAlaAlaGlnGluLeu 216
Db       771 GGGTCGGGCATGCAGGGGGTGTTGCTGAGCTCCAAGTACTTGAAGGCCGCACAGGAGCTT 830

Qy       217 LeuAspGluValValAsnIleValGlyLysSerIleLysGlyAspAspGlnLysLysAsp 236
Db       831 CTCGAAGAGGTTGTTAATGTTGGGAACACCGGAATCAAAGCTGAGATGCTGAAGAAGGCC 890

Qy       237 AsnSerMetAsnLysGluSerMetProLeuAlaSerAspValAsnThrAsnSerSerGly 256
Db       891 AGCGGCCAAAGTAAG---------CCGGGTGGAGAATCAGCGGCACTGAAGGAGGAAGGA 941

Qy       257 GlyGlyGluSerSerSerArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGln 276
Db       942 GGTGGCGACGGCAGTGGTAAGCGC-----------GGCGCAGAACTGTCTATGGCTGAG 989

Qy       277 ArgGlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeuGluGluValGluGln 296
Db       990 AGGCAGGAAATTCAGATGAAGAAGGCTAAGCTCATAAACATGCTTGATGAGGTTGAACAG 1049

Qy       297 ArgTyrArgGlnTyrHisHisGlnMetGlnIleIleValLeuSerPheGluGlnValAla 316
Db      1050 AGATACAGACAGTACCACAACCAGATGCAGATCGTGATATCCTCGTTTGAGCAAGCTGCG 1109

Qy       317 GlyIleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPhe 336
Db      1110 GGGATTGGCTCTGCAAGGACGTACACGGCCCTCGCGCTGCAGACCATCTCAAAGCAGTTC 1169

Qy       337 ArgCysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGlu 356
Db      1170 CGGTGCCTCAAGGACGCGATAGCAGGCCAGATTCGGGCCGCTAATAAGAGCTTGGGCGAG 1229

Qy       357 GluGluGly 359
```

```
Db      1230 GAGGATGGC 1238

RESULT 2
US-09-640-211A-77
; Sequence 77, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
;  APPLICANT: Wood, Marion
;  APPLICANT:  Shenk, Michael A.
;  APPLICANT:  McGrath, Annette
;  APPLICANT:  Glenn, Matthew
;  TITLE OF INVENTION: Compositions and Methods for the
;  TITLE OF INVENTION:  Modification of Gene Transcription
;  FILE REFERENCE: 11000.1021C1U
;  CURRENT APPLICATION NUMBER: US/09/640,211A
;  CURRENT FILING DATE:  2000-08-16
;  NUMBER OF SEQ ID NOS: 2368
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 77
;   LENGTH: 529
;   TYPE: DNA
;   ORGANISM: Eucalyptus grandis
US-09-640-211A-77

Alignment Scores:
Pred. No.:              1.76e-46        Length:         529
Score:                  493.00          Matches:        103
Percent Similarity:     69.15%          Conservative:   27
Best Local Similarity:  54.79%          Mismatches:     38
Query Match:            13.84%          Indels:         20
DB:                     4               Gaps:           5

US-10-624-201A-2 (1-688) x US-09-640-211A-77 (1-529)

Qy       355 GlyGluGluGluGlyLeuGlyGlyLysIleGluGlySerArgLeuLysPheValAspHis 374
Db         2 GGAAGCAGTAAAGGAGTGGGA-----------ATACCCCGGCTTCGTTTTTTGGACCAG 49

Qy       375 HisLeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGlnProAsnAlaTrpArg 394
Db        50 CAGCTCAGGCAGCAACGAGCCCTACAGCAGCTTGGGATGATGCAACAGCATGCATGGAGG 109

Qy       395 ProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeuPheGluHis 414
Db       110 CCGCAAAGAGGACTTCCTGAGAGTTCTGTTTCTATTCTTCGGGCCTGGCTATTTGAGCAT 169

Qy       415 PheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGlnThrGlyLeu 434
Db       170 TTTCTTCATCCTTACCCAAAGGATTCTGACAAAATCCTGCTTGCAAGGCAGACAGGCTTG 229

Qy       435 ThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLysProMet 454
Db       230 ACAAGAAGTCAGGTCTCGAATTGGTTCATCAATGCAAGAGTGCGTCTCTGGAAACCTATG 289

Qy       455 ValGluGluMetTyrLeuGluGluValLysAsnGlnGluGlnAsnSerThrAsnThrSer 474
Db       290 GTCGAAGAAATGTACAAAGAAGAGATTGGGGATGCGGAAATGGACTCCAACTCATCCTCC 349

Qy       475 GlyAspAsnLysAsnLysGluThrAsnIleSerAlaProAsnGluGluLysHisProIle 494
Db       350 GACACAGCCAAGCCAAAAACAGGAGATATCAAGTCCTCCATGGAGGACCGG--------- 400

Qy       495 IleThrSerSerLeuLeuGlnAspGlyIleThrThrThrGln------------AlaGlu 510
Db       401 ------GTGGAAGAAGTGCAACAGAGTTCAACAGCTACACAGAGATGCAGCTCAGGCCAG 454

Qy       511 IleSerThrSerThrIleSerThrSerPro---------ThrAlaGlyAlaSerLeuHis 527
Db       455 CTCATGGACTCATCATTCGACCGGACTCCAGATGTCGAAATGGCAGGC------------ 502

Qy       528 HisAlaHisAsnPheSerPheLeu 535
```

```
Db        503 CACTCTGTGGGATTCAACTACCTG 526

RESULT 3
US-09-640-211A-102
; Sequence 102, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
;  APPLICANT: Wood, Marion
;  APPLICANT:  Shenk, Michael A.
;  APPLICANT:  McGrath, Annette
;  APPLICANT:  Glenn, Matthew
;  TITLE OF INVENTION: Compositions and Methods for the
;  TITLE OF INVENTION:  Modification of Gene Transcription
;  FILE REFERENCE: 11000.1021C1U
;  CURRENT APPLICATION NUMBER: US/09/640,211A
;  CURRENT FILING DATE:  2000-08-16
;  NUMBER OF SEQ ID NOS: 2368
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 102
;   LENGTH: 381
;   TYPE: DNA
;   ORGANISM: Eucalyptus grandis
US-09-640-211A-102

Alignment Scores:
Pred. No.:                2.96e-41       Length:          381
Score:                    445.00         Matches:         84
Percent Similarity:       80.80%         Conservative:    17
Best Local Similarity:    67.20%         Mismatches:      20
Query Match:              12.49%         Indels:          4
DB:                       4              Gaps:            1

US-10-624-201A-2 (1-688) x US-09-640-211A-102 (1-381)

Qy        344 AlaGluGlnValLysAlaThrSerLysSerLeuGlyGluGluGluGlyLeuGlyGly--- 362
              |||   |||:::|||   :::   :::   |||||||||:::::::|||   |||
Db          3 GCCGCCCAACTGAAGCACAGCTGCGAGCTACTGGGTGAGAAGGACGGCGCGGGCAGCTCC 62

Qy        363 ---------LysIleGluGlySerArgLeuLysPheValAspHisHisLeuArgGlnGln 379
                      |||   |||       |||||||||   :::|||       ||||||||||||
Db         63 GGTATAACCAAGGGCGAGACACCACGGCTCAAGTTGCTCGACCAGAGCCTGAGGCAGCAG 122

Qy        380 ArgAlaLeuGlnGlnIleGlyMetMetGlnProAsnAlaTrpArgProGlnArgGlyLeu 399
              ||||||      |||:::|||||||:::       |||||||||||||||||||||||
Db        123 AGGGCTTTCCACCAGATGGGCATGATGGAGCAAGAGGCCTGGAGGCCGCAGCGGGGCCTG 182

Qy        400 ProGluArgAlaValSerValLeuArgAlaTrpLeuPheGluHisPheLeuHisProTyr 419
              |||||||||:::|||:::::::||||||||||||||||||||||||||||||||||||||
Db        183 CCGGAGCGGTCGGTCAACATACTGCGTGCATGGCTCTTCGAGCATTTCTTGCATCCGTAT 242

Qy        420 ProLysAspSerAspLysIleMetLeuAlaLysGlnThrGlyLeuThrArgSerGlnVal 439
              |||   |||:::|||||   :::||||||:::|||||||||||:::|||:::||||||
Db        243 CCAAGTGACGCTGATAAGCATCTGTTGGCTCGACAGACTGGTCTCTCCAGAAACCAGGTC 302

Qy        440 SerAsnTrpPheIleAsnAlaArgValArgLeuTrpLysProMetValGluGluMetTyr 459
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 TCGAATTGGTTCATAAATGCCAGGGTCCGGTTGTGGAAACCCATGGTGGAGGAGATGTAC 362

Qy        460 LeuGluGluValLys 464
                 :::|||   |||
Db        363 CAGCAAGAGTCCAAA 377

RESULT 4
US-09-640-211A-1472
; Sequence 1472, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
;  APPLICANT:. Wood, Marion
;  APPLICANT:  Shenk, Michael A.
;  APPLICANT:  McGrath, Annette
;  APPLICANT:  Glenn, Matthew
;  TITLE OF INVENTION: Compositions and Methods for the
;  TITLE OF INVENTION:  Modification of Gene Transcription
;  FILE REFERENCE: 11000.1021C1U
;  CURRENT APPLICATION NUMBER: US/09/640,211A
;  CURRENT FILING DATE:  2000-08-16
;  NUMBER OF SEQ ID NOS: 2368
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1472
;   LENGTH: 381
;   TYPE: DNA
;   ORGANISM: Eucalyptus grandis
US-09-640-211A-1472

Alignment Scores:
Pred. No.:                2.96e-41       Length:          381
Score:                    445.00         Matches:         84
Percent Similarity:       80.80%         Conservative:    17
Best Local Similarity:    67.20%         Mismatches:      20
Query Match:              12.49%         Indels:          4
DB:                       4              Gaps:            1

US-10-624-201A-2 (1-688) x US-09-640-211A-1472 (1-381)

Qy        344 AlaGluGlnValLysAlaThrSerLysSerLeuGlyGluGluGluGlyLeuGlyGly--- 362
              |||   |||:::|||   :::   :::   |||||||||:::::|||   |||
Db          3 GCCGCCCAACTGAAGCACAGCTGCGAGCTACTGGGTGAGAAGGACGGCGCGGGCAGCTCC 62

Qy        363 ---------LysIleGluGlySerArgLeuLysPheValAspHisHisLeuArgGlnGln 379
                      |||   |||       |||||||||   :::|||       ||||||||||||
Db         63 GGTATAACCAAGGGCGAGACACCACGGCTCAAGTTGCTCGACCAGAGCCTGAGGCAGCAG 122

Qy        380 ArgAlaLeuGlnGlnIleGlyMetMetGlnProAsnAlaTrpArgProGlnArgGlyLeu 399
              ||||||      |||:::|||||||:::       |||||||||||||||||||||||
Db        123 AGGGCTTTCCACCAGATGGGCATGATGGAGCAAGAGGCCTGGAGGCCGCAGCGGGGCCTG 182

Qy        400 ProGluArgAlaValSerValLeuArgAlaTrpLeuPheGluHisPheLeuHisProTyr 419
              |||||||||:::|||:::::::||||||||||||||||||||||||||||||||||||||
Db        183 CCGGAGCGGTCGGTCAACATACTGCGTGCATGGCTCTTCGAGCATTTCTTGCATCCGTAT 242

Qy        420 ProLysAspSerAspLysIleMetLeuAlaLysGlnThrGlyLeuThrArgSerGlnVal 439
              |||   |||:::|||||   :::||||||:::|||||||||||:::|||:::||||||
Db        243 CCAAGTGACGCTGATAAGCATCTGTTGGCTCGACAGACTGGTCTCTCCAGAAACCAGGTC 302

Qy        440 SerAsnTrpPheIleAsnAlaArgValArgLeuTrpLysProMetValGluGluMetTyr 459
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 TCGAATTGGTTCATAAATGCCAGGGTCCGGTTGTGGAAACCCATGGTGGAGGAGATGTAC 362

Qy        460 LeuGluGluValLys 464
                 :::|||   |||
Db        363 CAGCAAGAGTCCAAA 377

RESULT 5
US-09-640-211A-1854
; Sequence 1854, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
;  APPLICANT: Wood, Marion
;  APPLICANT:  Shenk, Michael A.
;  APPLICANT:  McGrath, Annette
;  APPLICANT:  Glenn, Matthew
;  TITLE OF INVENTION: Compositions and Methods for the
;  TITLE OF INVENTION:  Modification of Gene Transcription
;  FILE REFERENCE: 11000.1021C1U
;  CURRENT APPLICATION NUMBER: US/09/640,211A
;  CURRENT FILING DATE:  2000-08-16
;  NUMBER OF SEQ ID NOS: 2368
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1854
;   LENGTH: 316
;   TYPE: DNA
;   ORGANISM: Pinus radiata
US-09-640-211A-1854

Alignment Scores:
Pred. No.:                1.52e-38       Length:          316
```

```
Score:                 420.00        Matches:       79
Percent Similarity:    87.38%        Conservative:  11
Best Local Similarity: 76.70%        Mismatches:    11
Query Match:           11.79%        Indels:        2
DB:                    4             Gaps:          1

US-10-624-201A-2 (1-688) x US-09-640-211A-1854 (1-316)

Qy        380 ArgAlaLeuGlnGlnIleGlyMetMetGlnProAsnAlaTrpArgProGlnArgGlyLeu 399
              ||||||||||||||||:::||||||:::|||   :::||||||||||||||||||||||
Db          2 CGGGCTCTCCAACAATTAGGCATGATTCAGCAGCATGCTTGGAGGCCACAGAGAGGACTT 61

Qy        400 ProGluArgAlaValSerValLeuArgAlaTrpLeuPheGluHisPheLeuHisProTyr 419
              |||||||||:::||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 CCCGAGCGATCTGTTTCTGTCTTACGGGCTTGGCTATTTGAACATTTTCTTCATCCGTAT 121

Qy        420 ProLysAspSerAspLysIleMetLeuAlaLysGlnThrGlyLeuThrArgSerGlnVal 439
              |||||||||:::||||||   |||||||||:::||||||||||||||||||:::||||||
Db        122 CCAAAAGATGCAGACAAACATATGCTCGCGAGACAGACTGGGCTTACCAGAAATCAGGTC 181

Qy        440 SerAsnTrpPheIleAsnAlaArgValArgLeuTrpLysProMetValGluGluMetTyr 459
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 TCAAATTGGTTTATAAATGCACGTGTACGCCTCTGGAAGCCTATGGTGGAAGAGATGTAT 241

Qy        460 LeuGluGluValLysAsnGlnGluGlnAsnSerThrAsnThrSerGlyAspAsnLysAsn 479
              :::||||||   |||      |||   :::      :::   |||   |||
Db        242 GTGGAGGAAACAAAGGAGGCAGAAGTAGAC------CATGGATCAAATGATAAAACAGGT 295

Qy        480 LysGluThr 482
              |||||||:::
Db        296 AAGGAGAGT 304

RESULT 6
US-09-640-211A-359
; Sequence 359, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
;  APPLICANT: Wood, Marion
;  APPLICANT:  Shenk, Michael A.
;  APPLICANT:  McGrath, Annette
;  APPLICANT:  Glenn, Matthew
;  TITLE OF INVENTION: Compositions and Methods for the
;  TITLE OF INVENTION:  Modification of Gene Transcription
;  FILE REFERENCE: 11000.1021C1U
;  CURRENT APPLICATION NUMBER: US/09/640,211A
;  CURRENT FILING DATE:  2000-08-16
;  NUMBER OF SEQ ID NOS: 2368
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 359
;   LENGTH: 462
;   TYPE: DNA
;   ORGANISM: Pinus radiata
US-09-640-211A-359

Alignment Scores:
Pred. No.:             2.99e-38      Length:        462
Score:                 420.00        Matches:       79
Percent Similarity:    87.38%        Conservative:  11
Best Local Similarity: 76.70%        Mismatches:    11
Query Match:           11.79%        Indels:        2
DB:                    4             Gaps:          1

US-10-624-201A-2 (1-688) x US-09-640-211A-359 (1-462)

Qy        380 ArgAlaLeuGlnGlnIleGlyMetMetGlnProAsnAlaTrpArgProGlnArgGlyLeu 399
              ||||||||||||||||:::||||||:::|||   :::||||||||||||||||||||||
Db          2 CGGGCTCTCCAACAATTAGGCATGATTCAGCAGCATGCTTGGAGGCCACAGAGAGGACTT 61

Qy        400 ProGluArgAlaValSerValLeuArgAlaTrpLeuPheGluHisPheLeuHisProTyr 419
              |||||||||:::||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 CCCGAGCGATCTGTTTCTGTCTTACGGGCTTGGCTATTTGAACATTTTCTTCATCCGTAT 121

Qy        420 ProLysAspSerAspLysIleMetLeuAlaLysGlnThrGlyLeuThrArgSerGlnVal 439
              |||||||||:::||||||   |||||||||:::||||||||||||||||||:::||||||
Db        122 CCAAAAGATGCAGACAAACATATGCTCGCGAGACAGACTGGGCTTACCAGAAATCAGGTC 181

Qy        440 SerAsnTrpPheIleAsnAlaArgValArgLeuTrpLysProMetValGluGluMetTyr 459
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 TCAAATTGGTTTATAAATGCACGTGTACGCCTCTGGAAGCCTATGGTGGAAGAGATGTAT 241

Qy        460 LeuGluGluValLysAsnGlnGluGlnAsnSerThrAsnThrSerGlyAspAsnLysAsn 479
              :::||||||   |||      |||   :::      :::   |||   |||
Db        242 GTGGAGGAAACAAAGGAGGCAGAAGTAGAC------CATGGATCAAATGATAAAACAGGT 295

Qy        480 LysGluThr 482
              |||||||:::
Db        296 AAGGAGAGT 304

RESULT 7
US-09-640-211A-1578
; Sequence 1578, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
;  APPLICANT: Wood, Marion
;  APPLICANT:  Shenk, Michael A.
;  APPLICANT:  McGrath, Annette
;  APPLICANT:  Glenn, Matthew
;  TITLE OF INVENTION: Compositions and Methods for the
;  TITLE OF INVENTION:  Modification of Gene Transcription
;  FILE REFERENCE: 11000.1021C1U
;  CURRENT APPLICATION NUMBER: US/09/640,211A
;  CURRENT FILING DATE:  2000-08-16
;  NUMBER OF SEQ ID NOS: 2368
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1578
;   LENGTH: 343
;   TYPE: DNA
;   ORGANISM: Pinus radiata
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(343)
;   OTHER INFORMATION: n = A,T,C or G
US-09-640-211A-1578

Alignment Scores:
Pred. No.:             2.28e-38      Length:        343
Score:                 419.00        Matches:       77
Percent Similarity:    82.46%        Conservative:  17
Best Local Similarity: 67.54%        Mismatches:    18
Query Match:           11.76%        Indels:        2
DB:                    4             Gaps:          1

US-10-624-201A-2 (1-688) x US-09-640-211A-1578 (1-343)

Qy        365 GluGlySerArgLeuLysPheValAspHisHisLeuArgGlnGlnArgAlaLeuGlnGln 384
              |||      ||||||     ::::::     :::|||||||||||||||||
Db          1 GAAACACCAAGGTTGGGATNTCTAGAACGAAGCATACGACAACAGCGCGCATTTCACCAC 60

Qy        385 IleGlyMetMetGlnProAsnAlaTrpArgProGlnArgGlyLeuProGluArgAlaVal 404
              :::|||:::|||:::   :::   |||||||||||||||||||||||||:::|||
Db         61 TTAGGATTGATGGAGCAGCACCCTTGGCGACCGCAGAGAGGACTTCCTGAACGCTCTGTN 120

Qy        405 SerValLeuArgAlaTrpLeuPheGluHisPheLeuHisProTyrProLysAspSerAsp 424
              ||||||||||||||||||||||||||||||||||||||||||||   |||:::|||
Db        121 TCTGTTCTTCGTGCATGGTTGTTTGAGCATTTTCTGCACCCGTATCCAACTGATGCAGAT 180

Qy        425 LysIleMetLeuAlaLysGlnThrGlyLeuThrArgSerGlnValSerAsnTrpPheIle 444
              |||   :::|||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAGCATATATTGGCTAAGCAAACTGGCCTTACAAGAAGTCAGGTATCAAATTGGTTTATA 240

Qy        445 AsnAlaArgValArgLeuTrpLysProMetValGluGluMetTyrLeuGluGluValLys 464
              ||||||||||||||||||||||||||||||||||||||||||:::||||||:::|||
Db        241 AATGCCAGGGTTAGACTATGGAAGCCCATGGTGGAGGAGATGTACATGGAAGAACTCAAG 300

Qy        465 AsnGln------GluGlnAsnSerThrAsnThrSerGlyAsp 476
                  :::      :::|||   :::   |||:::      :::
```

```
Db        301 GAAGAAAAAGTGGACCAAGGTACACACAATTCTGAAGCTGAA 342

RESULT 8
US-09-640-211A-418
; Sequence 418, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
;  APPLICANT: Wood, Marion
;  APPLICANT:  Shenk, Michael A.
;  APPLICANT:  McGrath, Annette
;  APPLICANT:  Glenn, Matthew
;  TITLE OF INVENTION: Compositions and Methods for the
;  TITLE OF INVENTION:  Modification of Gene Transcription
;  FILE REFERENCE: 11000.1021C1U
;  CURRENT APPLICATION NUMBER: US/09/640,211A
;  CURRENT FILING DATE:  2000-08-16
;  NUMBER OF SEQ ID NOS: 2368
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 418
;   LENGTH: 323
;   TYPE: DNA
;   ORGANISM: Pinus radiata
US-09-640-211A-418

Alignment Scores:
Pred. No.:              1.68e-37        Length:         323
Score:                  411.00          Matches:        74
Percent Similarity:     85.05%          Conservative:   17
Best Local Similarity:  69.16%          Mismatches:     14
Query Match:            11.54%          Indels:         2
DB:                     4               Gaps:           1

US-10-624-201A-2 (1-688) x US-09-640-211A-418 (1-323)

Qy        372 ValAspHisHisLeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGlnProAsn 391
Db          2 CTAGAACGAAGCATACGACAACAGCGCGCATTTCACCACTTAGGATTGATGGAGCAGCAC 61

Qy        392 AlaTrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeu 411
Db         62 CCTTGGCGACCGCAGAGAGGACTTCCTGAACGCTCTGTTTCTGTTCTTCGTGCATGGTTG 121

Qy        412 PheGluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGln 431
Db        122 TTTGAGCATTTTCTGCACCCGTATCCAACTGATGCAGATAAGCATATATTGGCTAAGCAA 181

Qy        432 ThrGlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrp 451
Db        182 ACTGGCCTTACAAGAAGTCAGGTATCAAATTGGTTTATAAATGCCAGGGTTAGACTATGG 241

Qy        452 LysProMetValGluGluMetTyrLeuGluGluValLysAsnGln------GluGlnAsn 469
Db        242 AAGCCCATGGTGGAGGAGATGTACATGGAAGAACTCAAGGAAGAAAAAGTGGACCAAGGT 301

Qy        470 SerThrAsnThrSerGlyAsp 476
Db        302 ACACACAATTCTGAAGCTGAA 322

RESULT 9
US-09-640-211A-37
; Sequence 37, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
;  APPLICANT: Wood, Marion
;  APPLICANT:  Shenk, Michael A.
;  APPLICANT:  McGrath, Annette
;  APPLICANT:  Glenn, Matthew
;  TITLE OF INVENTION: Compositions and Methods for the
;  TITLE OF INVENTION:  Modification of Gene Transcription
;  FILE REFERENCE: 11000.1021C1U
;  CURRENT APPLICATION NUMBER: US/09/640,211A
;  CURRENT FILING DATE:  2000-08-16
;  NUMBER OF SEQ ID NOS: 2368
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 37
;   LENGTH: 698
;   TYPE: DNA
;   ORGANISM: Eucalyptus grandis
US-09-640-211A-37

Alignment Scores:
Pred. No.:              4.41e-31        Length:         698
Score:                  360.00          Matches:        74
Percent Similarity:     69.23%          Conservative:   34
Best Local Similarity:  47.44%          Mismatches:     46
Query Match:            10.10%          Indels:         2
DB:                     4               Gaps:           2

US-10-624-201A-2 (1-688) x US-09-640-211A-37 (1-698)

Qy        203 MetValLeuGlySerLysTyrLeuLysAlaAlaGlnGluLeuLeuAspGluValValAsn 222
Db        222 GTCATTCCCAATTCTAGGTACTTGAGGGTGGCGCAGCAACTGCTTGATGAAATAGTGAAT 281

Qy        223 IleValGlyLysSerIleLysGlyAspAspGlnLysLysAspAsnSerMetAsnLysGlu 242
Db        282 ---GTGCGAAAGGCTTTGAAGCGCTCTGATGACGCGAATGACCAATCTAGACAT---GAG 335

Qy        243 SerMetProLeuAlaSerAspValAsnThrAsnSerSerGlyGlyGlyGluSerSerSer 262
Db        336 AACCAAAGGAGCCCCAAAGATGCGGATGGGGGTTCCAAGAACGAAGCATCCTCAAATCCC 395

Qy        263 ArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGlnArgGlnGluLeuGlnMet 282
Db        396 CAAGAATCCGCCAGTAACTCTAGCGAGCTTTCTGCTGCTGAAAAACAAGATTTGCAGAAC 455

Qy        283 LysLysAlaLysLeuLeuAlaMetLeuGluGluValGluGlnArgTyrArgGlnTyrHis 302
Db        456 AAGCTCACGAAGCTCTTGTCCATGTTGGACGAGGTTGATAAAAGGTACAAGCAGTACTAT 515

Qy        303 HisGlnMetGlnIleIleValLeuSerPheGluGlnValAlaGlyIleGlySerAlaLys 322
Db        516 CACCAAATGCAGATCGTGGTACAGTCTTTTGATACAATAGCAGGAAGCGGTGCAGCCAAG 575

Qy        323 SerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArgCysLeuLysAspAla 342
Db        576 CCCTACACGGCGCTTGCGCTCCAGAGGATATCCCGCCACTTCCGGTGCCTGCATGACGCG 635

Qy        343 IleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGluGluGlu 358
Db        636 ATCACGGGTCAAATTCAAGCAACCCGTAAAAGTCTCGGAGAGCAAGAC 683

RESULT 10
US-09-640-211A-400
; Sequence 400, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
;  APPLICANT: Wood, Marion
;  APPLICANT:  Shenk, Michael A.
;  APPLICANT:  McGrath, Annette
;  APPLICANT:  Glenn, Matthew
;  TITLE OF INVENTION: Compositions and Methods for the
;  TITLE OF INVENTION:  Modification of Gene Transcription
;  FILE REFERENCE: 11000.1021C1U
;  CURRENT APPLICATION NUMBER: US/09/640,211A
;  CURRENT FILING DATE:  2000-08-16
;  NUMBER OF SEQ ID NOS: 2368
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 400
;   LENGTH: 690
;   TYPE: DNA
;   ORGANISM: Pinus radiata
US-09-640-211A-400

Alignment Scores:
Pred. No.:              1.7e-27         Length:         690
```

```
Score:                  328.50          Matches:        79
Percent Similarity:     61.64%          Conservative:   19
Best Local Similarity:  49.69%          Mismatches:     46
Query Match:            9.22%           Indels:         15
DB:                     4               Gaps:           5

US-10-624-201A-2 (1-688) x US-09-640-211A-400 (1-690)

Qy        181 SerSerSerSerProArgThrAsnAsnValThrIleArgGlyThr---LeuAspGlySer 199
              |||   :::   |||     |||             ||||||        ||:::
Db        214 AGTGTGAATGGCCCTGAGTGTAACAAACCGATGCATCATGGTACGAATGCAATTGGAGCA 273

Qy        200 Ser------SerAsnMetValLeuGlySerLysTyrLeuLysAlaAlaGlnGluLeuLeu 217
              :::      ||||||   :::   ||||||||||||   ||||||||||:::||||||
Db        274 GCAGAGCTCTCAAACACTTTAACTGGTTCCAAATATTTTAAAGCAGCACAGCAATTACTT 333

Qy        218 AspGluValValAsnIleValGlyLysSerIleLysGlyAspAspGlnLysLysAspAsn 237
              ||||||||||||||   |||||||||   ||||||   |||
Db        334 GATGAAGTTGTAAAT---GTTGGAAAGGGTATCAAGTCTGATTCAGTCAACCATCAAAAA 390

Qy        238 SerMetAsn---------------LysGluSerMetProLeuAlaSerAspValAsnThr 252
              |||                        :::         :::|||::::::      |||
Db        391 TCCCAAACATGGTTTGGTGCAATATCTGACAAAAAGAATATTGCAACTGAAGCTACTACA 450

Qy        253 AsnSer-----------------SerGlyGlyGlyGluSerSerSerArgGlnLysAsn 266
              |||                    :::|||         |||:::         ||||||
Db        451 AATGACCGAACAACATCTGCAATAACAGGAGCTTCAATTTCTGCAGAAGTAATGAAAAAC 510

Qy        267 GluValAlaValGluLeuThrThrAlaGlnArgGlnGluLeuGlnMetLysLysAlaLys 286
              |||   |||      ||||||   |||   ||||||||||||||||||||||||||||||
Db        511 GAGCATGCTTTTGGACTCACACCAGCTGATAGACAAGAACTTCAGATGAAAAAGGCAAAG 570

Qy        287 LeuLeuAlaMetLeuGluGluValGluGlnArgTyrArgGlnTyrHisHisGlnMetGln 306
              |||:::|||||||||:::||||||::::::|||||||||||||||:::||||||||||||
Db        571 CTTGTTGCCATGTTGGATGAGGTGGATCGAAGGTACAGACAGTACTATCATCAGATGCAA 630

Qy        307 IleIleValLeuSerPheGluGlnValAlaGlyIleGlySerAlaLysSerTyrThr 325
              |||:::|||   |||||||||      ||||||   |||:::||||||:::||||||
Db        631 ATCGTTGTTTCATCGTTTGAGACCGCAGCTGGATTTGGGGCTGCCAAGACATACACT 687

RESULT 11
US-09-640-211A-111
; Sequence 111, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
;  APPLICANT: Wood, Marion
;  APPLICANT:  Shenk, Michael A.
;  APPLICANT:  McGrath, Annette
;  APPLICANT:  Glenn, Matthew
;  TITLE OF INVENTION: Compositions and Methods for the
;  TITLE OF INVENTION:  Modification of Gene Transcription
;  FILE REFERENCE: 11000.1021C1U
;  CURRENT APPLICATION NUMBER: US/09/640,211A
;  CURRENT FILING DATE:  2000-08-16
;  NUMBER OF SEQ ID NOS: 2368
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 111
;   LENGTH: 380
;   TYPE: DNA
;   ORGANISM: Eucalyptus grandis
US-09-640-211A-111

Alignment Scores:
Pred. No.:              6.27e-27        Length:         380
Score:                  319.50          Matches:        69
Percent Similarity:     68.60%          Conservative:   14
Best Local Similarity:  57.02%          Mismatches:     23
Query Match:            8.97%           Indels:         15
DB:                     4               Gaps:           4

US-10-624-201A-2 (1-688) x US-09-640-211A-111 (1-380)

Qy        411 LeuPheGluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLys 430
              ||||||||||||||||||||||||||||||||||||||||||:::||||||||||||
Db          2 CTCTTCGAACACTTTCTCCACCCCTACCCGAAGGATTCGGACAAAGTCATGCTGGCCAAA 61

Qy        431 GlnThrGlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeu 450
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 CAGACAGGGCTCACTAGAAGCCAGGTGTCGAATTGGTTTATAAATGCTCGAGTTCGGCTT 121

Qy        451 TrpLysProMetValGluGluMetTyrLeuGluGluValLysAsnGlnGluGlnAsn--- 469
              ||||||||||||||||||||||||||   ||||||:::|||   ||||||||||||
Db        122 TGGAAGCCGATGGTGGAGGAGATGTACACGGAGGAAATCAAGGAGCAAGAACAGAATGGG 181

Qy        470 --------------------SerThrAsnThrSerGlyAspAsnLysAsnLysGluThr 482
                                  |||   :::         |||:::   :::|||   :::
Db        182 GGAGGAGCAGAGGAAAAACCAAGCAAGAGTGAACGCGAGGACTCAGCATCCAAGTCCTCT 241

Qy        483 AsnIle------SerAlaProAsnGluGluLysHisProIleIleThrSerSerLeuLeu 500
              :::         |||||||||      ::::::         |||   |||
Db        242 GGCCTCCAGGACAAGGCCCCCAACTCCAATGAGAACAGC------ACCAAGAGCTTCAAA 295

Qy        501 GlnAspGlyIleThrThrThrGlnAlaGlu---------IleSerThrSerThrIleSer 517
                    ||||||:::         :::         ||||||||::::::   |||
Db        296 CCAAAGGAGATCACCTCGAGGAACCACGACACCCCTGCCATCTCTACTAATTCGGCTTCC 355

Qy        518 Thr 518
              :::
Db        356 TCC 358

RESULT 12
US-09-640-211A-1246
; Sequence 1246, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
;  APPLICANT: Wood, Marion
;  APPLICANT:  Shenk, Michael A.
;  APPLICANT:  McGrath, Annette
;  APPLICANT:  Glenn, Matthew
;  TITLE OF INVENTION: Compositions and Methods for the
;  TITLE OF INVENTION:  Modification of Gene Transcription
;  FILE REFERENCE: 11000.1021C1U
;  CURRENT APPLICATION NUMBER: US/09/640,211A
;  CURRENT FILING DATE:  2000-08-16
;  NUMBER OF SEQ ID NOS: 2368
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1246
;   LENGTH: 380
;   TYPE: DNA
;   ORGANISM: Eucalyptus grandis
US-09-640-211A-1246

Alignment Scores:
Pred. No.:              6.27e-27        Length:         380
Score:                  319.50          Matches:        69
Percent Similarity:     68.60%          Conservative:   14
Best Local Similarity:  57.02%          Mismatches:     23
Query Match:            8.97%           Indels:         15
DB:                     4               Gaps:           4

US-10-624-201A-2 (1-688) x US-09-640-211A-1246 (1-380)

Qy        411 LeuPheGluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLys 430
              ||||||||||||||||||||||||||||||||||||||||||:::||||||||||||
Db          2 CTCTTCGAACACTTTCTCCACCCCTACCCGAAGGATTCGGACAAAGTCATGCTGGCCAAA 61

Qy        431 GlnThrGlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeu 450
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 CAGACAGGGCTCACTAGAAGCCAGGTGTCGAATTGGTTTATAAATGCTCGAGTTCGGCTT 121

Qy        451 TrpLysProMetValGluGluMetTyrLeuGluGluValLysAsnGlnGluGlnAsn--- 469
              ||||||||||||||||||||||||||   ||||||:::|||   ||||||||||||
Db        122 TGGAAGCCGATGGTGGAGGAGATGTACACGGAGGAAATCAAGGAGCAAGAACAGAATGGG 181

Qy        470 --------------------SerThrAsnThrSerGlyAspAsnLysAsnLysGluThr 482
                                  |||   :::         |||:::   :::|||   :::
```

```
Db        182 GGAGGAGCAGAGGAAAAACCAAGCAAGAGTGAACGCGAGGACTCAGCATCCAAGTCCTCT 241

Qy        483 AsnIle------SerAlaProAsnGluGluLysHisProIleIleThrSerSerLeuLeu 500
                 :::        |||||||||      ::::::          |||   |||
Db        242 GGCCTCCAGGACAAGGCCCCCAACTCCAATGAGAACAGC------ACCAAGAGCTTCAAA 295

Qy        501 GlnAspGlyIleThrThrThrGlnAlaGlu---------IleSerThrSerThrIleSer 517
                  ||||||:::        :::         |||||||||::::::   |||
Db        296 CCAAAGGAGATCACCTCGAGGAACCACGACACCCCTGCCATCTCTACTAATTCGGCTTCC 355

Qy        518 Thr 518
                 :::
Db        356 TCC 358

RESULT 13
US-09-640-211A-1241
; Sequence 1241, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
;  APPLICANT: Wood, Marion
;  APPLICANT:  Shenk, Michael A.
;  APPLICANT:  McGrath, Annette
;  APPLICANT:  Glenn, Matthew
;  TITLE OF INVENTION: Compositions and Methods for the
;  TITLE OF INVENTION:  Modification of Gene Transcription
;  FILE REFERENCE: 11000.1021C1U
;  CURRENT APPLICATION NUMBER: US/09/640,211A
;  CURRENT FILING DATE:  2000-08-16
;  NUMBER OF SEQ ID NOS: 2368
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1241
;   LENGTH: 366
;   TYPE: DNA
;   ORGANISM: Eucalyptus grandis
US-09-640-211A-1241

Alignment Scores:
Pred. No.:              2.17e-24        Length:         366
Score:                  297.00          Matches:        65
Percent Similarity:     64.62%          Conservative:   19
Best Local Similarity:  50.00%          Mismatches:     30
Query Match:            8.34%           Indels:         16
DB:                     4               Gaps:           4

US-10-624-201A-2 (1-688) x US-09-640-211A-1241 (1-366)

Qy        413 GluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGlnThr 432
              ||||||||||||||||||||||||||||||||||||||||||:::||||||:::||||||
Db          1 GAGCATTTTCTTCATCCTTACCCAAAGGATTCTGACAAAATCCTGCTTGCAAGGCAGACA 60

Qy        433 GlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLys 452
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCTTGACAAGAAGTCAGGTCTCGAATTGGTTCATCAATGCAAGAGTGCGTCTCTGGAAA 120

Qy        453 ProMetValGluGluMetTyrLeuGluGluValLysAsnGlnGluGlnAsnSerThrAsn 472
              ||||||||||||||||||||  ||||||:::  :::  |||  :::|||  :::
Db        121 CCTATGGTCGAAGAAATGTACAAAGAAGAGATTGGGGATGCGGAAATGGACTCCAACTCA 180

Qy        473 ThrSerGlyAspAsnLysAsnLysGluThrAsnIleSerAlaProAsnGluGluLysHis 492
              :::|||        |||  |||      :::|||  :::      |||::::::
Db        181 TCCTCCGACACAGCCAAGCCAAAAACAGGAGATATCAAGTCCTCCATGGAGGACCGG--- 237

Qy        493 ProIleIleThrSerSerLeuLeuGlnAspGlyIleThrThrThrGln------------ 508
                                  :::|||         |||  ||||||
Db        238 ------------GTGGAAGAAGTGCAACAGAGTTCAACAGCTACACAGAGATGCAGCTCA 285

Qy        509 AlaGluIleSerThrSerThrIleSerThrSerPro---------ThrAlaGlyAlaSer 525
              ::::::     |||:::      :::|||              ||||||
Db        286 GGCCAGCTCATGGACTCATCATTCGACCGGACTCCAGATGTCGAAATGGCAGGC------ 339

Qy        526 LeuHisHisAlaHisAsnPheSerPheLeu 535
                   |||:::     |||::::::|||
Db        340 ------CACTCTGTGGGATTCAACTACCTG 363

RESULT 14
US-09-640-211A-1845
; Sequence 1845, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
;  APPLICANT: Wood, Marion
;  APPLICANT:  Shenk, Michael A.
;  APPLICANT:  McGrath, Annette
;  APPLICANT:  Glenn, Matthew
;  TITLE OF INVENTION: Compositions and Methods for the
;  TITLE OF INVENTION:  Modification of Gene Transcription
;  FILE REFERENCE: 11000.1021C1U
;  CURRENT APPLICATION NUMBER: US/09/640,211A
;  CURRENT FILING DATE:  2000-08-16
;  NUMBER OF SEQ ID NOS: 2368
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1845
;   LENGTH: 171
;   TYPE: DNA
;   ORGANISM: Pinus radiata
US-09-640-211A-1845

Alignment Scores:
Pred. No.:              9.38e-21        Length:         171
Score:                  260.00          Matches:        47
Percent Similarity:     92.86%          Conservative:   5
Best Local Similarity:  83.93%          Mismatches:     4
Query Match:            7.30%           Indels:         0
DB:                     4               Gaps:           0

US-10-624-201A-2 (1-688) x US-09-640-211A-1845 (1-171)

Qy        367 SerArgLeuLysPheValAspHisHisLeuArgGlnGlnArgAlaLeuGlnGlnIleGly 386
              |||||||||  |||||||||      ||||||||||||||||||||||||||:::|||
Db          4 TCCCGTCTTCACTTTGTTGATCAACAATTACGACAACAGCGAGCTCTTCAGCAGCTAGGA 63

Qy        387 MetMetGlnProAsnAlaTrpArgProGlnArgGlyLeuProGluArgAlaValSerVal 406
              |||:::|||   :::||||||||||||||||||||||||||||||||||||||||:::
Db         64 ATGATACAGCAGCATGCCTGGAGACCACAAAGAGGGCTTCCAGAGAGGGCCGTTTCTATT 123

Qy        407 LeuArgAlaTrpLeuPheGluHisPheLeuHisProTyrProLysAsp 422
              |||||||||||||||||||||||||||||||||||||||||||:::
Db        124 CTCCGGGCTTGGCTATTTGAGCATTTCCTTCATCCGTACCCCAAAAAT 171

RESULT 15
US-09-640-211A-351
; Sequence 351, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
;  APPLICANT: Wood, Marion
;  APPLICANT:  Shenk, Michael A.
;  APPLICANT:  McGrath, Annette
;  APPLICANT:  Glenn, Matthew
;  TITLE OF INVENTION: Compositions and Methods for the
;  TITLE OF INVENTION:  Modification of Gene Transcription
;  FILE REFERENCE: 11000.1021C1U
;  CURRENT APPLICATION NUMBER: US/09/640,211A
;  CURRENT FILING DATE:  2000-08-16
;  NUMBER OF SEQ ID NOS: 2368
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 351
;   LENGTH: 260
;   TYPE: DNA
;   ORGANISM: Pinus radiata
US-09-640-211A-351

Alignment Scores:
Pred. No.:              5.66e-20        Length:         260
Score:                  256.00          Matches:        54
Percent Similarity:     79.49%          Conservative:   8
Best Local Similarity:  69.23%          Mismatches:     7
Query Match:            7.18%           Indels:         9
```

```
DB:                    4             Gaps:          3

US-10-624-201A-2 (1-688) x US-09-640-211A-351 (1-260)

Qy        413 GluHisPheLeu-HisProTyrProLysAspSerAspLysIleMetLeuAlaLysGlnTh 432
              :::|||||||||  ||||||||||||||||||:::||||||   :::||||||:::|||||
Db          9 CAGCATTTCCTTCCATCCGTACCCAAAAGATGCTGACAAACATTTACTAGCAAGACAGAC 68

Qy        432 rGlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLy 452
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         69 TGGACTGACCAGAAGCCAGGTTTCAAATTGGTTTATAAATGCACGTGTCCGCCTTTGGAA 128

Qy        452 sProMetValGluGluMetTyrLeuGluGluValLysAsnGlnGlu---GlnAsnSer-- 470
              ||||||||||||||||||||:::||||||:::::      |||   ||||||
Db        129 ACCCATGGTGGAAGAAATGTATATGGAGGAACTTAGAGAGGCCGAAACACAGAATCATGC 188

Qy        471 ----------------ThrAsnThrSerGlyAspAsnLysAsnLysGluThr 482
                           |||      ||||||   |||   |||:::||||||
Db        189 AGCAGATTCGAAGGTAACAACAGAAAGTGGTCAAAAC---AATGAAGAAACG 237


Search completed: July 29, 2005, 18:59:21
Job time : 316 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 29, 2005, 14:48:42 ; Search time 1102 Seconds
(without alignments)
4040.818 Million cell updates/sec

Title: US-10-624-201A-2
Perfect score: 3563
Sequence: 1 MYYQGTSDNTNIQADHQQRH.....GNKRFPTQLLPDFVTGNLGT 688

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 7287783 seqs, 3236178273 residues

Total number of hits satisfying chosen parameters: 14575566

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10624201/runat_27072005_122959_10961/app_query.fasta_1.839
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10624201_@CGN_1_1_789_@runat_27072005_122959_10961
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:*
22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:*
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1353	38.0	3033	18	US-10-424-599-130628	Sequence 130628,
2	1174.5	33.0	1555	18	US-10-425-114-13353	Sequence 13353, A
3	1150.5	32.3	2043	9	US-09-938-842A-1337	Sequence 1337, Ap
4	1150.5	32.3	2043	11	US-09-938-842A-1337	Sequence 1337, Ap
5	1150.5	32.3	2043	21	US-10-495-918-153	Sequence 153, App
6	1148.5	32.2	2385	17	US-10-225-066A-335	Sequence 335, App
7	1148.5	32.2	2385	17	US-10-374-780A-2671	Sequence 2671, Ap
8	1148.5	32.2	2385	22	US-10-225-066A-335	Sequence 335, App
9	1079.5	30.3	1822	18	US-10-424-599-133558	Sequence 133558,
10	1078	30.3	1606	18	US-10-425-114-7804	Sequence 7804, Ap
11	1050.5	29.5	3029	19	US-10-437-963-60154	Sequence 60154, A
12	1037	29.1	1892	18	US-10-424-599-63874	Sequence 63874, A
13	999	28.0	2313	18	US-10-425-114-15195	Sequence 15195, A
14	999	28.0	2970	20	US-10-425-115-97705	Sequence 97705, A
15	984	27.6	1716	18	US-10-425-114-9882	Sequence 9882, Ap
16	984	27.6	3311	18	US-10-424-599-141176	Sequence 141176,
17	945	26.5	3028	19	US-10-437-963-41007	Sequence 41007, A
18	941.5	26.4	1498	18	US-10-425-114-33974	Sequence 33974, A
19	926.5	26.0	2799	19	US-10-437-963-31585	Sequence 31585, A
20	912.5	25.6	2943	20	US-10-425-115-118586	Sequence 118586,
21	864	24.2	1649	18	US-10-425-114-20970	Sequence 20970, A
c 22	845.5	23.7	3135	19	US-10-437-963-67929	Sequence 67929, A
23	824	23.1	1884	9	US-09-938-842A-905	Sequence 905, App
24	824	23.1	1884	11	US-09-938-842A-905	Sequence 905, App
25	824	23.1	1983	14	US-10-286-264-17	Sequence 17, Appl
26	824	23.1	1983	17	US-10-225-066A-819	Sequence 819, App
27	824	23.1	1983	17	US-10-374-780A-2481	Sequence 2481, Ap
28	824	23.1	1983	18	US-10-412-699B-637	Sequence 637, App
29	824	23.1	1983	22	US-10-225-066A-819	Sequence 819, App
30	823.5	23.1	1905	19	US-10-437-963-89736	Sequence 89736, A
31	818.5	23.0	1645	18	US-10-424-599-65908	Sequence 65908, A
32	810.5	22.7	1637	18	US-10-425-114-13359	Sequence 13359, A
33	802	22.5	1599	9	US-09-938-842A-2154	Sequence 2154, Ap
34	802	22.5	1599	11	US-09-938-842A-2154	Sequence 2154, Ap
35	797.5	22.4	1727	18	US-10-425-114-34958	Sequence 34958, A
36	797.5	22.4	2545	17	US-10-225-068-235	Sequence 235, App
37	797.5	22.4	2545	17	US-10-374-780A-2329	Sequence 2329, Ap
38	797.5	22.4	2545	18	US-10-412-699B-267	Sequence 267, App
39	797.5	22.4	2545	21	US-10-225-068-235	Sequence 235, App
40	795	22.3	2486	18	US-10-424-599-55280	Sequence 55280, A
41	793.5	22.3	1736	20	US-10-425-115-55285	Sequence 55285, A
42	784	22.0	2736	18	US-10-424-599-37574	Sequence 37574, A
43	776.5	21.8	831	19	US-10-767-701-10301	Sequence 10301, A
44	775.5	21.8	1005	18	US-10-425-114-22835	Sequence 22835, A
45	775.5	21.8	2442	18	US-10-424-599-139178	Sequence 139178,

ALIGNMENTS

RESULT 1
US-10-424-599-130628
; Sequence 130628, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 130628

```
;   LENGTH: 3033
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_88966C.1
US-10-424-599-130628

Alignment Scores:
Pred. No.:              2.36e-124      Length:         3033
Score:                  1353.00        Matches:        350
Percent Similarity:     56.87%         Conservative:   93
Best Local Similarity:  44.93%         Mismatches:     160
Query Match:            37.97%         Indels:         176
DB:                     18             Gaps:           36

US-10-624-201A-2 (1-688) x US-10-424-599-130628 (1-3033)

Qy         16 HisGlnGlnArgHisAsnHisGlyAsnSerAsnAsnAsnAsn---------------- 29
              |||   :::|||:::||||||                  |||
Db        245 CACTGAGAGAGATACAATCATAGAATTGAAGAAGAACGCAACATGGCGACGTACTTTAAT 304

Qy         30 ------------------------------------IleGlnThrLeuTyrLeuMet 36
                                                   :::||||||||||||||||||
Db        305 GGAAGCACTTCGGAAATTCAACCTTCTGCAGATGGCGGCTTGCAAACGCTTTACCTTATG 364

Qy         37 AsnProAsnAsnTyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGlnGlnGlnLeu 56
              ||||||      |||:::              :::|||
Db        365 AACCCA---ACTTACGTA---------CCCTACGCTGACGCGCCGCACCACCCAACAACG 412

Qy         57 LeuPheLeuAsnSerSerProAlaAlaSerAsnAlaLeuCysHisAlaAsnIleGlnHis 76
              |||   :::|||      |||   |||   |||            |||||||||   |||
Db        413 CTCCTCGTCAAT------CCCAACGCACCCAACCTC---------GCGAATATCCCACAC 457

Qy         77 AlaProLeuGlnGlnGlnHisPheValGlyValProLeuProAlaValSerLeuHisAsp 96
              |||                                      |||   ||||||   :::
Db        458 GCG-----------------------------------CCGCCAGTGTCACCAAACCAA 481

Qy         97 GlnIleAsnHis------HisGlyLeuLeuGlnArgMet------TrpAsnAsnGlnAsp 112
              |||   :::|||      ||||||:::          :::       |||:::   |||
Db        482 CAACATCACCACGTGATTCACGGCGTCACAAGCATCATCGGATCCGGCAATTCCGATGAC 541

Qy        113 GlnSerGlnGlnValIleVal----------------------------ProSerSer 122
              |||:::     ::::::
Db        542 CACTCCCGACCGTCGTTGATCGGCGAAAACATCGCCGCCTTCCACGGCTTCTCCGGCGGG 601

Qy        123 ThrGlyValSerAlaThrSerCysGlyGlyIleThrThrAspLeuAlaSerGlnLeuAla 142
                 |||   ::::::|||:::          :::      :::|||          :::
Db        602 GCAGGCACTGCCTCCACCACCGCCCCTCGCCTTCACTATAACCTCTGGGGTCCGGTGGTT 661

Qy        143 PheGlnArgProIleProThrProGln--------------HisArgGlnGlnGlnGln 157
              |||         |||:::                     |||:::      |||
Db        662 GATCAGCCGGGAACGCCGTCTAGCAGCTCCGGTGGAGGCGGTTTCCGGCGGCCGAGTCAG 721

Qy        158 GlnGlnGlyGlyLeuSerLeuSerLeuSerProGlnLeuGlnGlnGlnIleSerPheAsn 177
              |||      ||||||||||||||||||||   |||
Db        722 CAG------GGTTTGTCTCTCAGTCTTTCATCACAG----------------------- 751

Qy        178 AsnAsnIleSerSerSerSerProArgThrAsnAsnValThrIleArgGlyThrLeuAsp 197
                                  :::||||||      ::::::   |||   ||||||
Db        752 -----------------------CAAACAAATTTTAGATCAGTTTCCGGCGAGCTTGAC 787

Qy        198 ----------------------------GlySerSer---------------------- 200
                                         |||:::|||
Db        788 GTGGCCGGACAGGGCCACGTTGCCGGAATTGGGAACTCGCCGATGTCGGCGTCGATCGGA 847

Qy        201 ---SerAsnMetValLeuGlySerLysTyrLeuLysAlaAlaGlnGluLeuLeuAspGlu 219
                 |||   :::::::::||||||||||||||||||||||||||||||||||||||||||
Db        848 GTTTCAGGTGTTATTATGGGGTCAAAGTACTTGAAAGCTGCACAGGAGCTTCTGGATGAA 907

Qy        220 ValValAsnIleValGlyLysSerIleLysGlyAspAspGlnLysLysAspAsn---Ser 238
              |||||||||   |||||||||   |||      :::::::::      ::::::
Db        908 GTTGTGAAC---GTGGGAAAAGGAATCTACAAAGAGGAGAAATTCTCAGAGAGCGTGAAA 964

Qy        239 MetAsnLysGluSer------MetProLeuAlaSerAspValAsnThrAsnSerSerGly 256
                 |||:::||||||                |||   |||              |||||||||
Db        965 GCGAATAGGGAATCAACTAATTCTGGTGCTGCTGGTGATGGTGGAGATGGTTCAAGTGGT 1024

Qy        257 GlyGlyGluSerSerSerArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGln 276
              |||||||||:::|||:::   ::::::              |||||||||:::|||||||||
Db       1025 GGAGGAGAAAACAGTGCAGGGAAACAA---------GTGGTTGAACTCAGCACCGCTCAG 1075

Qy        277 ArgGlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeuGluGluValGluGln 296
              |||||||||||||||||||||||:::|||||:::   ||||||:::|||||||||||||
Db       1076 AGGCAAGAGCTTCAGATGAAGAAGTCCAAGCTTGTGACCATGCTCGATGAGGTAGAACAA 1135

Qy        297 ArgTyrArgGlnTyrHisHisGlnMetGlnIleIleValLeuSerPheGluGlnValAla 316
              |||||||||||||||||||||||||||||||:::|||   ||||||||||||   |||
Db       1136 AGGTACAGACAATATCACCACCAAATGCAAATTGTGGTATCGTCATTTGAGCAAGCAGCA 1195

Qy        317 GlyIleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPhe 336
              |||   |||:::|||||||||||||||   |||||||||      ||||||||||||||
Db       1196 GGTTATGGGGCGGCAAAATCTTACACTGCCCTTGCCCTAAAAACTATCTCAAAGCAATTC 1255

Qy        337 ArgCysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGlu 356
              |||||||||||||||||||:::   |||:::|||||||||||||||:::|||||||||
Db       1256 AGGTGCCTGAAAGATGCAATCTCTGCACAAATCAAAGCCACGAGCAAGACATTGGGTGAA 1315

Qy        357 GluGluGlyLeuGlyGlyLysIleGluGlySerArgLeuLysPheValAspHisHisLeu 376
              ::::::   ||||||   |||:::|||||||||||||:::::|||||||||||||||||
Db       1316 GATGATTGCTTAGGAGTTAAAGTAGAAGGGTCAAGGCTTAGGTATGTTGACCACCATCTC 1375

Qy        377 ArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGlnProAsnAlaTrpArgProGln 396
              |||||||||||||||||||||:::||||||:::|||||||||||||||||||||||||||
Db       1376 AGGCAACAAAGGGCACTTCAGCAGCTAGGAATGATTCAACCCAATGCATGGAGGCCCCAA 1435

Qy        397 ArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeuPheGluHisPheLeu 416
              |||||||||||||||||||||||||||:::|||||||||||||||||||||||||||||
Db       1436 AGAGGCTTGCCTGAACGAGCTGTTTCCATTCTTCGGGCTTGGCTTTTTGAGCATTTCCTT 1495

Qy        417 HisProTyrProLysAspSerAspLysIleMetLeuAlaLysGlnThrGlyLeuThrArg 436
              |||||||||||||||||||||||||||:::|||||||||||||||||||||||   |||
Db       1496 CACCCTTATCCAAAGGACTCCGATAAAGTTATGCTTGCTAAACAAACTGGACTTGCTAGG 1555

Qy        437 SerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLysProMetValGlu 456
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1556 AGCCAGGTGTCAAACTGGTTTATCAATGCCCGAGTTCGGCTTTGGAAGCCAATGGTTGAA 1615

Qy        457 GluMetTyrLeuGluGluValLysAsnGlnGluGln------AsnSerThrAsnThrSer 474
              |||||||||||||||||:::|||   ||||||      |||   ||||||
Db       1616 GAAATGTACTTGGAAGAAATCAAGGAGCATGAACAAGGTAATGGCTCAGAAAACACCAAA 1675

Qy        475 GlyAspAsnLysAsnLysGlu------------ThrAsnIleSerAlaProAsn---Glu 489
                           :::||||||          |||::::::      :::   :::
Db       1676 TCAAAAGAATCGAGCAAGGAGTTGGCTTCAACAGCAAATGTTGCACTCGATCATCTCCAA 1735

Qy        490 GluLysHisProIleIleThrSerSerLeuGluGlnAspGlyIleThrThrThrGlnAla 509
              ||||||            :::                        |||:::
Db       1736 TCAAAACATGAAAGCTTCAACAACCAA-----------------AATACTTCCCCAACC 1777

Qy        510 GluIleSerThrSerThrIleSerThrSerProThrAlaGlyAlaSerLeuHisHisAla 529
              |||||||||||||:::   |||   ||||||      |||:::        |||:::
Db       1778 GAGATCTCAACAAGCTCT---TCAATGTCTCCAATGGGAGGGTCCCTTCAATCCCACTCT 1834

Qy        530 HisAsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrThrValAspHis 549
                    |||         ||||||   :::|||
Db       1835 ---GGTTTCCATCTTGCAGGATCATCTGACATG----------------------CAA 1867

Qy        550 IleGluAsnAsnAlaLysLysGlnArgAsnAspMetHisLysPheSerProSerSerIle 569
              |||   :::      |||   :::   ::::::|||         ||||||||||||||
Db       1868 ATTAGGAGTCCAAACAAACCAAGAAGTTCTGAAATGCAAAAC---TCCCCAAGTAGTATC 1924

Qy        570 LeuSerSerValAspMetGluAlaLys--------------------AlaArgGluSer 582
              |||   ||||||||||||||   |||                        |||::::::
Db       1925 CTT---TCAGTGGACATGGAAATGAAGCATAGTGGTGACCATGGAAATAATAGAGATGCC 1981

Qy        583 SerAsnLys--------------------------------GlyPheThrAsnProLeu 591
```

```
Db       1982 AACACAAAGTTTGGTATTGAAAGGCATCACCAACAAAAAGATGGGTAC------CCTTTG 2035

Qy        592 Met----------------------------AlaAlaTyrAlaMetGlyAspPheGly 601

Db       2036 ATGACTAGCAATCCAAACCATGGTGGGGGGTTTGGAGCATTCACCATGGAAGACATTGGA 2095

Qy        602 ---ArgPheAspProHisAspGlnGlnMetThrAlaAsnPheHisGlyAsnAsnGlyVal 620

Db       2096 AGCAGATTTAATGTCACTACTGAGCAACTAGCTTCAAGGTTCCATGGA---AACGGTGTT 2152

Qy        621 SerLeuThrLeuGlyLeuProProSerGluAsnLeuAlaMetProValSerGlnGlnAsn 640

Db       2153 TCTCTCACACTTGGACTTCCCCACAATGAGAACCTTTCTATGTCAGGAACTCAACAACAT 2212

Qy        641 ---TyrLeuSerAsnAsp------LeuGlySerArgSerGluMetGlySerHisTyrAsn 657

Db       2213 GGGTTCCTCTCACAGAATATACACCTGGGAAGGAGACTTGAAATGGGAACAAATGGAAAT 2272

Qy        658 ArgMet-----------------------------------------GlyTyrGluAsn 663

Db       2273 GAGTTTTGTACTATCAACACTGCTCCACCTTCTTCTCACTCAGGAACCACTTATGAGAGC 2332

Qy        664 IleAspPheGlnSerGlyAsnLysArgProThrGlnLeuLeuProAspPheVal 682

Db       2333 ATTGATATTCAAAACAGAAAGAGGTTTGTTGCTCATCAGTTATTACCAGATTTCGTG 2389

RESULT 2
US-10-425-114-13353
; Sequence 13353, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 13353
;   LENGTH: 1555
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: JC-GMST02400042B10_FLI
US-10-425-114-13353

Alignment Scores:
Pred. No.:              6.33e-107      Length:          1555
Score:                  1174.50        Matches:          271
Percent Similarity:     66.06%         Conservative:      58
Best Local Similarity:  54.42%         Mismatches:        90
Query Match:            32.96%         Indels:            79
DB:                     18             Gaps:              20

US-10-624-201A-2 (1-688) x US-10-425-114-13353 (1-1555)

Qy        240 AsnLysGluSer------MetProLeuAlaSerAspValAsnThrAsnSerSerGlyGly 257

Db          5 AATAGGGAATCAACTAATTCTGGTGCTGCTGGTGATGGTGGAGATGGTTCAAGTGGTGGA 64

Qy        258 GlyGluSerSerSerArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGlnArg 277

Db         65 GGAGAAAACAGTGCAGGGAAACAA---------GTGGTTGAACTCAGCACCGCTCAGAGG 115

Qy        278 GlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeuGluGluValGluGlnArg 297

Db        116 CAAGAGCTTCAGATGAAGAAGTCCAAGCTTGTGACCATGCTCGATGAGGTAGAACAAAGG 175

Qy        298 TyrArgGlnTyrHisHisGlnMetGlnIleIleValLeuSerPheGluGlnValAlaGly 317

Db        176 TACAGACAATATCACCACCAAATGCAAATTGTGGTATCGTCATTTGAGCAGGCAGCAGGT 235

Qy        318 IleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArg 337

Db        236 TATGGGGCGGCAAAATCTTACACTGCCCTTGCCCTAAAGACTATCTCAAAGCAATTCAGG 295

Qy        338 CysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGluGlu 357

Db        296 TGCCTTAAAGATGCAATCTCTGCACAAATCAAAGCCACGAGCAAGACATTGGGTGAAGAT 355

Qy        358 GluGlyLeuGlyGlyLysIleGluGlySerArgLeuLysPheValAspHisHisLeuArg 377

Db        356 GATTGTTTAGGAGTTAAAGTAGAAGGGTCGAGGCTTAGGTTTGTTGACCACCATCTCAGG 415

Qy        378 GlnGlnArgAlaLeuGlnGlnIleGlyMetMetGlnProAsnAlaTrpArgProGlnArg 397

Db        416 CAACAAAGGGCACTTCAGCAGCTAGGAATGATTCAACCCAATGCATGGAGGCCCCAAAGA 475

Qy        398 GlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeuPheGluHisPheLeuHis 417

Db        476 GGCTTGCCTGAACGAGCTGTTTCCATTCTTCGGGCTTGGCTTTTTGAGCATTTCCTTCAC 535

Qy        418 ProTyrProLysAspSerAspLysIleMetLeuAlaLysGlnThrGlyLeuThrArgSer 437

Db        536 CCTTATCCAAAGGACTCCGATAAAGTTATGCTTGCTAAACAAACTGGACTTGCTAGGAGC 595

Qy        438 GlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLysProMetValGluGlu 457

Db        596 CAGGTGTCAAACTGGTTTATCAATGCCCGAGTTCGGCTTTGGAAGCCAATGGTTGAAGAA 655

Qy        458 MetTyrLeuGluGluValLysAsnGlnGluGln------AsnSerThrAsnThrSerGly 475

Db        656 ATGTACTTGGAAGAAATCAAGGAGCATGAACAAGGTAATGGCTCAGAAAACACCAAATCA 715

Qy        476 AspAsnLysAsnLysGlu------------ThrAsnIleSerAlaProAsn---GluGlu 490

Db        716 AAAGAATCGAGCAAGGAGTTGGCTTCAACAGCAAATGTTGCACTCGATCATCTCCAATCA 775

Qy        491 LysHisProIleIleThrSerSerLeuLeuGlnAspGlyIleThrThrThrGlnAlaGlu 510

Db        776 AAACATGAAAGCTTCAACAACCAA-----------------AATACTTCCCCAACCGAG 817

Qy        511 IleSerThrSerThrIleSerThrSerProThrAlaGlyAlaSerLeuHisHisAlaHis 530

Db        818 ATCTCAACAAGCTCT---TCAATGTCTCCAATGGGAGGGTCCCTTCAATCCCACTCT--- 871

Qy        531 AsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrThrValAspHisIle 550

Db        872 GGTTTCCATCTTGCAGGATCATCTGACATG----------------------CAAATT 907

Qy        551 GluAsnAsnAlaLysLysGlnArgAsnAspMetHisLysPheSerProSerSerIleLeu 570

Db        908 AGGAGTCCAAACAAACCAAGAAGTTCTGAAATGCAAAAC---TCCCCAAGTAGTATCCTT 964

Qy        571 SerSerValAspMetGluAlaLys--------------------AlaArgGluSerSer 583

Db        965 ---TCAGTGGACATGGAAATGAAGCATAGTGGTGACCATGGAAATAATAGAGATGCCAAC 1021

Qy        584 AsnLys--------------------------------GlyPheThrAsnProLeuMet 592

Db       1022 ACAAAGTTTGGTATTGAAAGGCATCACCAACAAAAAGATGGGTAC------CCTTTGATG 1075

Qy        593 ----------------------------AlaAlaTyrAlaMetGlyAspPheGly--- 601

Db       1076 ACTAGCAATCCAAACCATGGTGGGGGGTTTGGAGCATTCACCATGGAAGACATTGGAAGC 1135

Qy        602 ArgPheAspProHisAspGlnGlnMetThrAlaAsnPheHisGlyAsnAsnGlyValSer 621

Db       1136 AGATTTAATGTCACTACTGAGCAACTAGCTTCAAGGTTCCATGGA---AACGGTGTTTCT 1192
```

```
Qy       622 LeuThrLeuGlyLeuProProSerGluAsnLeuAlaMetProValSerGlnGlnAsn--- 640
Db      1193 CTCACACTTGGACTTCCCCACAATGAGAACCTTTCTATGTCAGGAACTCAACAACATGGG 1252

Qy       641 TyrLeuSerAsnAsp------LeuGlySerArgSerGluMetGlySerHisTyrAsnArg 658
Db      1253 TTCCTCTCACAGAATATACACCTGGGAAGGAGACTTGAAATGGGAACAAATGGAAATGAG 1312

Qy       659 Met----------------------------------------GlyTyrGluAsnIle 664
Db      1313 TTTTGTACTATCAACACTGCTCCACCTTCTTCTCACTCAGGAACCACTTATGAGAGCATT 1372

Qy       665 AspPheGlnSerGlyAsnLysArgPheProThrGlnLeuLeuProAspPheVal 682
Db      1373 GATATTCAAAACAGAAAGAGGTTTGTTGCTCATCAGTTATTACCAGATTTCGTG 1426
```

RESULT 3
US-09-938-842A-1337
; Sequence 1337, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIP1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1337
; LENGTH: 2043
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1337

Alignment Scores:

Pred. No.:	2.4e-104	Length:	2043
Score:	1150.50	Matches:	306
Percent Similarity:	51.81%	Conservative:	81
Best Local Similarity:	40.96%	Mismatches:	201
Query Match:	32.29%	Indels:	159
DB:	9	Gaps:	28

US-10-624-201A-2 (1-688) x US-09-938-842A-1337 (1-2043)

```
Qy        22 HisGlyAsn-----------------SerAsnAsnAsnAsnIleGlnThrLeuTyrLeu 35
Db        16 CACGGAAACCCACCGGAGATCTCTGCCGGATCCGACGGTGGTCTTCAAACGTTGATCCTC 75

Qy        36 MetAsnProAsnAsnTyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGlnGlnGln 55
Db        76 ATGAATCCAACTACTTACGTTCAG---TACACCCAACAAGACAACGACTCGAACAACAAC 132

Qy        56 LeuLeuPheLeuAsnSerSerProAlaAlaSerAsnAlaLeuCysHisAlaAsnIle--- 74
Db       133 -----------AACAACAGCAACAATAGCAACAACAACAACACAAACACAAACACAAAC 180

Qy        75 ----------------------------GlnHisAlaProLeu------GlnGlnGln 82
Db       181 AACAACAACAGTAGTTTCGTTTTCCTCGATTCCCACGCGCCGCAGCCAAACGCGAGCCAG 240

Qy        83 HisPheValGlyValProLeu--------------ProAlaValSerLeuHisAspGln 97
Db       241 CAGTTCGTCGGAATACCACTCTCAGGTCACGAAGCTGCTTCCATTACAGCCGCCGACAAC 300

Qy        98 IleAsn---HisHisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnSerGlnGln 116
Db       301 ATCTCCGTACTTCACGGTTATCCTCCGCGCGTGCAGTACAGTCTCTACGGTAGCCACCAA 360

Qy       117 ValIleValProSerSerThrGlyValSerAlaThrSerCysGlyGlyIleThrThrAsp 136
Db       361 GTG---------------------------------------------------------- 363

Qy       137 LeuAlaSerGlnLeuAlaPheGlnArgProIleProThrProGlnHisArgGln------ 154
Db       364 ----------------------------------GATCCCACTCACCAGCAAGCCGCG 387

Qy       155 ------GlnGlnGlnGlnGlnGlyGlyLeuSerLeuSerLeuSerProGlnLeuGlnGln 172
Db       388 TGTGAGACGCCACGCGCGCAGCAAGGCCTCTCTTTAACCCTCTCGTCTCAACAGCAGCAG 447

Qy       173 GlnIleSerPheAsnAsn--------------AsnIleSerSerSerSerProArgThr 187
Db       448 CAACAGCAACATCATCAACAACACCAGCCTATTCACGTCGGATTCGGGTCCGGACATGGA 507

Qy       188 AsnAsnValThrIleArgGlyThrLeuAspGlySerSer--------------SerAsn 202
Db       508 GAAGATATCCGGGTCGGGTCTGGCTCTACAGGATCGGGGGTAACAAACGGTATAGCTAAT 567

Qy       203 MetValLeuGlySerLysTyrLeuLysAlaAlaGlnGluLeuLeuAspGluValValAsn 222
Db       568 CTTGTT---AGCTCCAAGTACTTGAAGGCAGCACAAGAGCTTCTTGACGAAGTAGTCAAC 624

Qy       223 IleValGlyLysSerIleLysGlyAspAspGln------------LysLysAspAsnSer 238
Db       625 GCTGATTCCGATGACATGAACGCTAAATCCCAACTATTCTCATCGAAAAAGGGTAGTTGC 684

Qy       239 MetAsnLysGluSerMetProLeuAlaSer---AspValAsnThrAsnSerSerGlyGly 257
Db       685 GGAAATGATAAACCTGTCGGAGAATCATCGGCCGGCGCTGGAGGAGAAGGTTCCGGTGGC 744

Qy       258 GlyGluSerSerSerArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGlnArg 277
Db       745 GGAGCAGAAGCAGCCGGGAAACGT---------CCGGTGGAGCTAGGCACGGCAGAGAGA 795

Qy       278 GlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeuGluGluValGluGlnArg 297
Db       796 CAAGAAATACAGATGAAGAAAGCAAAACTTAGTAACATGCTTCATGAGGTGGAGCAGAGA 855

Qy       298 TyrArgGlnTyrHisHisGlnMetGlnIleIleValLeuSerPheGluGlnValAlaGly 317
Db       856 TATAGACAGTACCACCAGCAGATGCAGATGGTGATCTCTTCGTTCGAGCAAGCGGCAGGG 915

Qy       318 IleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArg 337
Db       916 ATAGGATCAGCGAAGTCATACACGTCGCTAGCATTGAAAACCATATCAAGACAGTTCCGT 975

Qy       338 CysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGluGlu 357
Db       976 TGCTTGAAAGAGGCGATCGCTGGTCAGATAAAAGCGGCCAACAAGAGTCTTGGGGAGGAA 1035

Qy       358 GluGlyLeuGly------GlyLysIleGluGlySerArgLeuLysPheValAspHisHis 375
Db      1036 GATTCAGTGTCTGGTGTTGGGAGGTTTGAGGGGTCGAGGCTCAAGTTCGTGGACCACCAC 1095

Qy       376 LeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGln---------ProAsnAla 392
Db      1096 TTGAGACAGCAAAGAGCTCTTCAACAACTGGGAATGATTCAACATCCTTCCAATAATGCT 1155

Qy       393 TrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeuPhe 412
Db      1156 TGGAGACCTCAACGTGGTCTCCCAGAACGAGCCGTCTCAGTTCTCCGTGCTTGGCTCTTC 1215

Qy       413 GluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGlnThr 432
Db      1216 GAACACTTTCTTCATCCATACCCTAAGGATTCGGACAAGCACATGCTAGCTAAGCAAACA 1275

Qy       433 GlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLys 452
```

```
Db       1276 GGACTCACTCGTAGCCAGGTGTCGAACTGGTTTATAAACGCGAGAGTTCGGTTATGGAAA 1335

Qy        453 ProMetValGluGluMetTyrLeuGluGluValLysAsnGlnGluGlnAsnSerThrAsn 472

Db       1336 CCAATGGTGGAGGAGATGTACATGGAGGAAATGAAGGAGCAGGCAAAGAACATG------ 1389

Qy        473 ThrSerGlyAspAsnLysAsnLysGluThrAsnIleSerAlaProAsnGluGluLysHis 492

Db       1390 -----------GGATCCATGGAAAAGACTCCTTTGGATCAAAGCAACGAAGAT------ 1431

Qy        493 ProIleIleThrSerSerLeuLeuGlnAspGlyIleThrThrThrGlnAlaGluIleSer 512

Db       1432 ----------------------------------------TCTGCTTCAAAGTCA 1446

Qy        513 ThrSerThrIleSerThrSerProThrAlaGlyAlaSerLeuHis------HisAlaHis 530

Db       1447 ACAAGTAACCAAGAAAAGAGCCCAATGGCGGACACTAATTACCATATGAATCCCAATCAC 1506

Qy        531 AsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrThrValAspHisIle 550

Db       1507 AACGGTGACCTAGAAGGCGTCACTGGAATGCAAGGAAGCCCCAAGAGACTAAGAACCAGC 1566

Qy        551 GluAsnAsnAlaLysLysGlnArgAsnAspMetHisLysPheSerProSerSerIleLeu 570

Db       1567 GACGAGACAATGATGCAGCCAATAAAT------GCGGATTTCAGCTCCAACGAGAAGCTC 1620

Qy        571 SerSerValAspMetGluAlaLysAlaArgGluSerSerAsnLysGlyPheThrAsnPro 590

Db       1621 ACGATGAAGATTCTAGAAGAACGGCAAGGGATAAGATCAGATGGTGGCTAC------CCT 1674

Qy        591 LeuMet---------AlaAlaTyrAlaMetGlyAspPheGlyArgPheAsp---ProHis 606

Db       1675 TTCATGGGTAATTTCGGGCAATACCAAATGGATGAGATGTCAAGATTTGATGTAGTCTCA 1734

Qy        607 AspGlnGlnMetThrAlaAsnPheHis------GlyAsnAsnGlyValSerLeuThrLeu 624

Db       1735 GACCAGGAGCTCATGGCGCAAAGGTACTCAGGAAACAACAATGGCGTGTCCCTCACGTTA 1794

Qy        625 GlyLeuProProSerGluAsnLeuAlaMetProValSerGln---------GlnAsnTyr 641

Db       1795 GGTTTACCTCATTGTGATAGCTTGTCGTCCACGCACCATCAGGGTTTCATGCAGACCCAC 1854

Qy        642 LeuSerAsnAspLeuGlySerArgSerGluMetGly---------------------- 653

Db       1855 CATGGGATTCCTATAGGGAGAAGAGTGAAAATAGGAGAAACAGAGGAATATGGACCCGCC 1914

Qy        654 ----------------------------------------------------SerHis 655

Db       1915 ACCATCAATGGTGGTAGCTCGACCACAACCGCACATTCATCAGCGGCAGCTGCCGCGGCT 1974

Qy        656 TyrAsnArgMetGlyTyrGluAsnIleAspPheGlnSerGlyAsnLysArgPheProThr 675

Db       1975 TACAATGGGATGAACATACAGAAC-----------------CAGAAGAGATATGTGGCT 2016

Qy        676 GlnLeuLeuProAspPheVal 682

Db       2017 CAGTTATTGCCCGACTTCGTT 2037

RESULT 4
US-09-938-842A-1337
; Sequence 1337, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
;  APPLICANT: Harper, Jeff
;  APPLICANT:  Kreps, Joel
;  APPLICANT:  Wang, Xun
;  APPLICANT:  Zhu, Tong
;  TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
;  TITLE OF INVENTION:  SAME, AND METHODS OF USE
;  FILE REFERENCE: SCRIP1300-3
;  CURRENT APPLICATION NUMBER: US/09/938,842A
;  CURRENT FILING DATE:  2001-08-24
;  PRIOR APPLICATION NUMBER: US 60/227,866
;  PRIOR FILING DATE: 2000-08-24
;  PRIOR APPLICATION NUMBER: US 60/264,647
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/300,111
;  PRIOR FILING DATE: 2001-06-22
;  NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1337
;   LENGTH: 2043
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-938-842A-1337

Alignment Scores:
Pred. No.:              2.4e-104      Length:          2043
Score:                  1150.50       Matches:         306
Percent Similarity:     51.81%        Conservative:    81
Best Local Similarity:  40.96%        Mismatches:      201
Query Match:            32.29%        Indels:          159
DB:                     11            Gaps:            28

US-10-624-201A-2 (1-688) x US-09-938-842A-1337 (1-2043)

Qy         22 HisGlyAsn-----------------SerAsnAsnAsnAsnIleGlnThrLeuTyrLeu 35

Db         16 CACGGAAACCCACCGGAGATCTCTGCCGGATCCGACGGTGGTCTTCAAACGTTGATCCTC 75

Qy         36 MetAsnProAsnAsnTyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGlnGlnGln 55

Db         76 ATGAATCCAACTACTTACGTTCAG---TACACCCAACAAGACAACGACTCGAACAACAAC 132

Qy         56 LeuLeuPheLeuAsnSerSerProAlaAlaSerAsnAlaLeuCysHisAlaAsnIle--- 74

Db        133 ------------AACAACAGCAACAATAGCAACAACAACAACACAAACACAAACACAAAC 180

Qy         75 ----------------------------GlnHisAlaProLeu------GlnGlnGln 82

Db        181 AACAACAACAGTAGTTTCGTTTTCCTCGATTCCCACGCGCCGCAGCCAAACGCGAGCCAG 240

Qy         83 HisPheValGlyValProLeu--------------ProAlaValSerLeuHisAspGln 97

Db        241 CAGTTCGTCGGAATACCACTCTCAGGTCACGAAGCTGCTTCCATTACAGCCGCCGACAAC 300

Qy         98 IleAsn---HisHisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnSerGlnGln 116

Db        301 ATCTCCGTACTTCACGGTTATCCTCCGCGCGTGCAGTACAGTCTCTACGGTAGCCACCAA 360

Qy        117 ValIleValProSerSerThrGlyValSerAlaThrSerCysGlyGlyIleThrThrAsp 136

Db        361 GTG---------------------------------------------------------- 363

Qy        137 LeuAlaSerGlnLeuAlaPheGlnArgProIleProThrProGlnHisArgGln------ 154

Db        364 ----------------------------------GATCCCACTCACCAGCAAGCCGCG 387

Qy        155 ------GlnGlnGlnGlnGlnGlyGlyLeuSerLeuSerLeuSerProGlnLeuGlnGln 172

Db        388 TGTGAGACGCCACGCGCGCAGCAAGGCCTCTCTTTAACCCTCTCGTCTCAACAGCAGCAG 447

Qy        173 GlnIleSerPheAsnAsn--------------AsnIleSerSerSerSerProArgThr 187

Db        448 CAACAGCAACATCATCAACAACACCAGCCTATTCACGTCGGATTCGGGTCCGGACATGGA 507

Qy        188 AsnAsnValThrIleArgGlyThrLeuAspGlySerSer--------------SerAsn 202

Db        508 GAAGATATCCGGGTCGGGTCTGGCTCTACAGGATCGGGGGTAACAAACGGTATAGCTAAT 567

Qy        203 MetValLeuGlySerLysTyrLeuLysAlaAlaGlnGluLeuLeuAspGluValValAsn 222

Db        568 CTTGTT---AGCTCCAAGTACTTGAAGGCAGCACAAGAGCTTCTTGACGAAGTAGTCAAC 624

Qy        223 IleValGlyLysSerIleLysGlyAspAspGln------------LysLysAspAsnSer 238
```

```
Db        625 GCTGATTCCGATGACATGAACGCTAAATCCCAACTATTCTCATCGAAAAAGGGTAGTTGC 684

Qy        239 MetAsnLysGluSerMetProLeuAlaSer---AspValAsnThrAsnSerSerGlyGly 257

Db        685 GGAAATGATAAACCTGTCGGAGAATCATCGGCCGGCGCTGGAGGAGAAGGTTCCGGTGGC 744

Qy        258 GlyGluSerSerSerArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGlnArg 277

Db        745 GGAGCAGAAGCAGCCGGGAAACGT---------CCGGTGGAGCTAGGCACGGCAGAGAGA 795

Qy        278 GlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeuGluGluValGluGlnArg 297

Db        796 CAAGAAATACAGATGAAGAAAGCAAAACTTAGTAACATGCTTCATGAGGTGGAGCAGAGA 855

Qy        298 TyrArgGlnTyrHisHisGlnMetGlnIleIleValLeuSerPheGluGlnValAlaGly 317

Db        856 TATAGACAGTACCACCAGCAGATGCAGATGGTGATCTCTTCGTTCGAGCAAGCGGCAGGG 915

Qy        318 IleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArg 337

Db        916 ATAGGATCAGCGAAGTCATACACGTCGCTAGCATTGAAAACCATATCAAGACAGTTCCGT 975

Qy        338 CysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGluGlu 357

Db        976 TGCTTGAAAGAGGCGATCGCTGGTCAGATAAAAGCGGCCAACAAGAGTCTTGGGGAGGAA 1035

Qy        358 GluGlyLeuGly------GlyLysIleGluGlySerArgLeuLysPheValAspHisHis 375

Db       1036 GATTCAGTGTCTGGTGTTGGGAGGTTTGAGGGGTCGAGGCTCAAGTTCGTGGACCACCAC 1095

Qy        376 LeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGln---------ProAsnAla 392

Db       1096 TTGAGACAGCAAAGAGCTCTTCAACAACTGGGAATGATTCAACATCCTTCCAATAATGCT 1155

Qy        393 TrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeuPhe 412

Db       1156 TGGAGACCTCAACGTGGTCTCCCAGAACGAGCCGTCTCAGTTCTCCGTGCTTGGCTCTTC 1215

Qy        413 GluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGlnThr 432

Db       1216 GAACACTTTCTTCATCCATACCCTAAGGATTCGGACAAGCACATGCTAGCTAAGCAAACA 1275

Qy        433 GlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLys 452

Db       1276 GGACTCACTCGTAGCCAGGTGTCGAACTGGTTTATAAACGCGAGAGTTCGGTTATGGAAA 1335

Qy        453 ProMetValGluGluMetTyrLeuGluGluValLysAsnGlnGluGlnAsnSerThrAsn 472

Db       1336 CCAATGGTGGAGGAGATGTACATGGAGGAAATGAAGGAGCAGGCAAAGAACATG------ 1389

Qy        473 ThrSerGlyAspAsnLysAsnLysGluThrAsnIleSerAlaProAsnGluGluLysHis 492

Db       1390 ------------GGATCCATGGAAAAGACTCCTTTGGATCAAAGCAACGAAGAT------ 1431

Qy        493 ProIleIleThrSerSerLeuLeuGlnAspGlyIleThrThrThrGlnAlaGluIleSer 512

Db       1432 ----------------------------------------------TCTGCTTCAAAGTCA 1446

Qy        513 ThrSerThrIleSerThrSerProThrAlaGlyAlaSerLeuHis------HisAlaHis 530

Db       1447 ACAAGTAACCAAGAAAAGAGCCCAATGGCGGACACTAATTACCATATGAATCCCAATCAC 1506

Qy        531 AsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrThrValAspHisIle 550

Db       1507 AACGGTGACCTAGAAGGCGTCACTGGAATGCAAGGAAGCCCCAAGAGACTAAGAACCAGC 1566

Qy        551 GluAsnAsnAlaLysLysGlnArgAsnAspMetHisLysPheSerProSerSerIleLeu 570

Db       1567 GACGAGACAATGATGCAGCCAATAAAT------GCGGATTTCAGCTCCAACGAGAAGCTC 1620

Qy        571 SerSerValAspMetGluAlaLysAlaArgGluSerSerAsnLysGlyPheThrAsnPro 590

Db       1621 ACGATGAAGATTCTAGAAGAACGGCAAGGGATAAGATCAGATGGTGGCTAC------CCT 1674

Qy        591 LeuMet---------AlaAlaTyrAlaMetGlyAspPheGlyArgPheAsp---ProHis 606

Db       1675 TTCATGGGTAATTTCGGGCAATACCAAATGGATGAGATGTCAAGATTTGATGTAGTCTCA 1734

Qy        607 AspGlnGlnMetThrAlaAsnPheHis------GlyAsnAsnGlyValSerLeuThrLeu 624

Db       1735 GACCAGGAGCTCATGGCGCAAAGGTACTCAGGAAACAACAATGGCGTGTCCCTCACGTTA 1794

Qy        625 GlyLeuProProSerGluAsnLeuAlaMetProValSerGln---------GlnAsnTyr 641

Db       1795 GGTTTACCTCATTGTGATAGCTTGTCGTCCACGCACCATCAGGGTTTCATGCAGACCCAC 1854

Qy        642 LeuSerAsnAspLeuGlySerArgSerGluMetGly---------------------- 653

Db       1855 CATGGGATTCCTATAGGGAGAAGAGTGAAAATAGGAGAAACAGAGGAATATGGACCCGCC 1914

Qy        654 ----------------------------------------------------SerHis 655

Db       1915 ACCATCAATGGTGGTAGCTCGACCACAACCGCACATTCATCAGCGGCAGCTGCCGCGGCT 1974

Qy        656 TyrAsnArgMetGlyTyrGluAsnIleAspPheGlnSerGlyAsnLysArgProThr 675

Db       1975 TACAATGGGATGAACATACAGAAC-----------------CAGAAGAGATATGTGGCT 2016

Qy        676 GlnLeuLeuProAspPheVal 682

Db       2017 CAGTTATTGCCCGACTTCGTT 2037
```

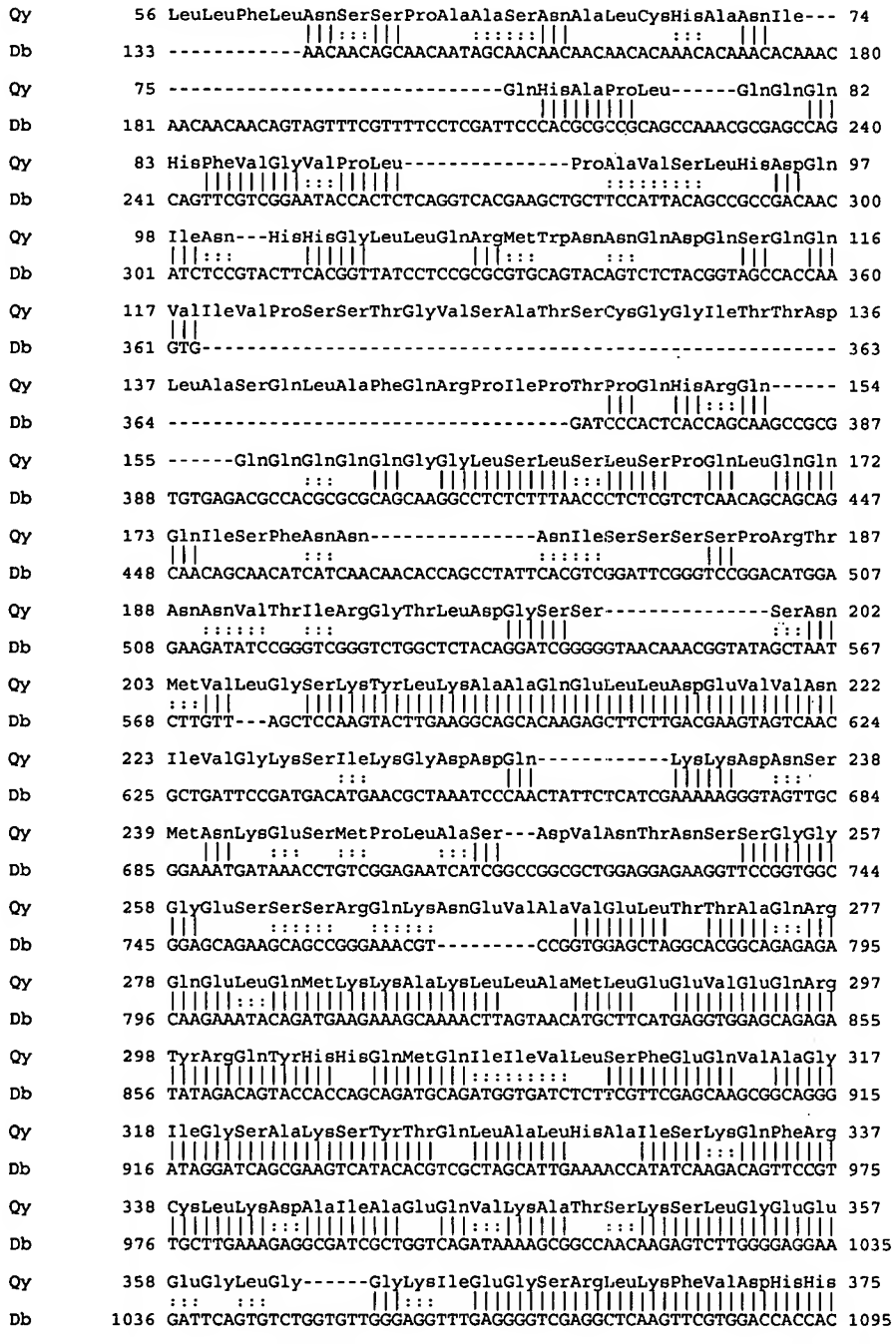
RESULT 5
US-10-495-918-153
; Sequence 153, Application US/10495918
; Publication No. US20050009187A1
; GENERAL INFORMATION:
; APPLICANT: RIKEN
; TITLE OF INVENTION: A STRESS-RESPONSIVE PROMOTER AND A GENE CORDING FOR A STRESS-RESPO
; TITLE OF INVENTION: TRANSCRIPTION FACTOR
; FILE REFERENCE: RFH14-024T
; CURRENT APPLICATION NUMBER: US/10/495,918
; CURRENT FILING DATE: 2004-05-18
; PRIOR APPLICATION NUMBER: JP 2001-353038
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: JP 2002-20329
; PRIOR FILING DATE: 2002-01-29
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 153
; LENGTH: 2043
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2040)
US-10-495-918-153

Alignment Scores:

Pred. No.:	2.4e-104	Length:	2043
Score:	1150.50	Matches:	306
Percent Similarity:	51.81%	Conservative:	81
Best Local Similarity:	40.96%	Mismatches:	201
Query Match:	32.29%	Indels:	159
DB:	21	Gaps:	28

US-10-624-201A-2 (1-688) x US-10-495-918-153 (1-2043)

```
Qy         22 HisGlyAsn-----------------SerAsnAsnAsnAsnIleGlnThrLeuTyrLeu 35

Db         16 CACGGAAACCCACCGGAGATCTCTGCCGGATCCGACGGTGGTCTTCAAACGTTGATCCTC 75

Qy         36 MetAsnProAsnAsnTyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGlnGlnGln 55

Db         76 ATGAATCCAACTACTTACGTTCAG---TACACCCAACAAGACAACGACTCGAACAACAAC 132
```



RESULT 6
US-10-225-066A-335

```
; Sequence 335, Application US/10225066A
; Publication No. US20030226173A1
; GENERAL INFORMATION:
;  APPLICANT: Mendel Biotechnology, Inc.
;  APPLICANT:  RATCLIFFE, Oliver
;  APPLICANT:  RIECHMANN, Jose Luis
;  APPLICANT:  ADAM, Luc J
;  APPLICANT:  DUBELL, Arnold T
;  APPLICANT:  HEARD, Jacqueline E
;  APPLICANT:  PILGRIM, Marsha L
;  APPLICANT:  JIANG, Cai-Zhong
;  APPLICANT:  REUBER, T. Lynne
;  APPLICANT:  CREELMAN, Robert A
;  APPLICANT:  PINEDA, Omaira
;  APPLICANT:  YU, Guo-Liang
;  APPLICANT:  BROUN, Pierre E
;  TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants
;  FILE REFERENCE: MBI0036-2 US
;  CURRENT APPLICATION NUMBER: US/10/225,066A
;  CURRENT FILING DATE:  2002-08-09
;  PRIOR APPLICATION NUMBER: 09/837,444
;  PRIOR FILING DATE: 2001-04-18
;  PRIOR APPLICATION NUMBER: 60/310,847
;  PRIOR FILING DATE: 2001-08-09
;  PRIOR APPLICATION NUMBER: 60/336,049
;  PRIOR FILING DATE: 2001-12-05
;  PRIOR APPLICATION NUMBER: 60/338,692
;  PRIOR FILING DATE: 2001-12-11
;  PRIOR APPLICATION NUMBER: 10/171,468
;  PRIOR FILING DATE: 2002-06-14
;  NUMBER OF SEQ ID NOS: 1122
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 335
;   LENGTH: 2385
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-10-225-066A-335

Alignment Scores:
Pred. No.:              4.81e-104      Length:          2385
Score:                  1148.50        Matches:         306
Percent Similarity:     51.67%         Conservative:    80
Best Local Similarity:  40.96%         Mismatches:      202
Query Match:            32.23%         Indels:          159
DB:                     17             Gaps:            28

US-10-624-201A-2 (1-688) x US-10-225-066A-335 (1-2385)

Qy         22 HisGlyAsn-----------------SerAsnAsnAsnAsnIleGlnThrLeuTyrLeu 35
              |||||||||                 ::::::      :::|||||||||   |||
Db        194 CACGGAAACCCACCGGAGATCTCTGCCGGATCCGACGGTGGTCTTCAAACGTTGATCCTC 253

Qy         36 MetAsnProAsnAsnTyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGlnGlnGln 55
              |||||||||      |||:::|||   ||||||      |||
Db        254 ATGAATCCAACTACTTACGTTCAG---TACACCCAACAAGACAACGACTCGAACAACAAC 310

Qy         56 LeuLeuPheLeuAsnSerSerProAlaAlaSerAsnAlaLeuCysHisAlaAsnIle--- 74
                    |||:::|||      ::::::|||        :::   |||
Db        311 ------------AACAACAGCAACAATAGCAACAACAACAACACAAACACAAACACAAAC 358

Qy         75 ----------------------------GlnHisAlaProLeu------GlnGlnGln 82
                                          |||||||||               |||
Db        359 AACAACAACAGTAGTTTCGTTTTCCTCGATTCCCACGCGCCGCAGCCAAACGCGAGCCAG 418

Qy         83 HisPheValGlyValProLeu---------------ProAlaValSerLeuHisAspGln 97
              |||||||||:::||||||                 :::::::::      |||
Db        419 CAGTTCGTCGGAATACCACTCTCAGGTCACGAAGCTGCTTCCATTACAGCCGCCGACAAC 478

Qy         98 IleAsn---HisHisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnSerGlnGln 116
              |||:::     ||||||      |||:::   :::         |||   |||
Db        479 ATCTCCGTACTTCACGGTTATCCTCCGCGCGTGCAGTACAGTCTCTACGGTAGCCACCAA 538

Qy        117 ValIleValProSerSerThrGlyValSerAlaThrSerCysGlyGlyIleThrThrAsp 136
                 |||
Db        539 GTG---------------------------------------------------------- 541

Qy        137 LeuAlaSerGlnLeuAlaPheGlnArgProIleProThrProGlnHisArgGln------ 154
                                                  |||   |||:::|||
Db        542 ----------------------------------GATCCCACTCACCAGCAAGCCGCG 565

Qy        155 ------GlnGlnGlnGlnGlnGlyGlyLeuSerLeuSerLeuSerProGlnLeuGlnGln 172
                     :::   |||   ||||||||||||:::||||||   |||   ||||||
Db        566 TGTGAGACGCCACGCGCGCAGCAAGGCCTCTCTTTAACCCTCTCGTCTCAACAGCAGCAG 625

Qy        173 GlnIleSerPheAsnAsn--------------AsnIleSerSerSerSerProArgThr 187
              |||         :::               ::::::       |||
Db        626 CAACAGCAACATCATCAACAACACCAGCCTATTCACGTCGGATTCGGGTCCGGACATGGA 685

Qy        188 AsnAsnValThrIleArgGlyThrLeuAspGlySerSer--------------SerAsn 202
                 ::::::   :::               ||||||                 :::|||
Db        686 GAAGATATCCGGGTCGGGTCTGGCTCTACAGGATCGGGGGTAACAAACGGTATAGCTAAT 745

Qy        203 MetValLeuGlySerLysTyrLeuLysAlaAlaGlnGluLeuLeuAspGluValValAsn 222
              :::|||      ||||||||||||||||||||||||||||||||||||||||||||||
Db        746 CTTGTT---AGCTCCAAGTACTTGAAGGCAGCACAAGAGCTTCTTGACGAAGTAGTCAAC 802

Qy        223 IleValGlyLysSerIleLysGlyAspAspGln-----------LysLysAspAsnSer 238
                        :::         |||         ||||||   :::
Db        803 GCTGATTCCGATGACATGAACGCTAAATCCCAACTATTCTCATCGAAAAAGGGTAGTTGC 862

Qy        239 MetAsnLysGluSerMetProLeuAlaSer---AspValAsnThrAsnSerSerGlyGly 257
                 |||   :::   :::      :::|||                 |||||||||
Db        863 GGAAATGATAAACCTGTCGGAGAATCATCGGCCGGCGCTGGAGGAGAAGGTTCCGGTGGC 922

Qy        258 GlyGluSerSerSerArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGlnArg 277
              |||       ::::::   ::::::             |||||||||   ||||||:::|||
Db        923 GGAGCAGAAGCAGCCGGGAAACGT---------CCGGTGGAGCTAGGCACGGCAGAGAGA 973

Qy        278 GlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeuGluGluValGluGlnArg 297
              ||||||:::||||||||||||||||||||      ||||||   ||||||||||||||||
Db        974 CAAGAAATACAGATGAAGAAAGCAAAACTTAGTAACATGCTTCATGAGGTGGAGCAGAGA 1033

Qy        298 TyrArgGlnTyrHisHisGlnMetGlnIleIleValLeuSerPheGluGlnValAlaGly 317
              ||||||||||||||   |||||||||::::::::   ||||||||||||   ||||||
Db       1034 TATAGACAGTACCACCAGCAGATGCAGATGGTGATCTCTTCGTTCGAGCAAGCGGCAGGG 1093

Qy        318 IleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArg 337
              ||||||||||||||||||||||||   |||||||||      ||||||:::|||||||||
Db       1094 ATAGGATCAGCGAAGTCATACACGTCGCTAGCATTGAAAACCATATCAAGACAGTTCCGT 1153

Qy        338 CysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGluGlu 357
              |||||||||:::|||||||||   |||:::||||||   :::|||||||||||||||||
Db       1154 TGCTTGAAAGAGGCGATCGCTGGTCAGATAAAAGCGGCCAACAAGAGTCTTGGGGAGGAA 1213

Qy        358 GluGlyLeuGly------GlyLysIleGluGlySerArgLeuLysPheValAspHisHis 375
              :::   :::         |||:::   ||||||||||||||||||||||||||||||||
Db       1214 GATTCAGTGTCTGGTGTTGGGAGGTTTGAGGGGTCGAGGCTCAAGTTCGTGGACCACCAC 1273

Qy        376 LeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGln---------ProAsnAla 392
              ||||||||||||||||||||||||||:::||||||:::|||            ||||||
Db       1274 TTGAGACAGCAAAGAGCTCTTCAACAACTGGGAATGATTCAACATCCTTCCAATAATGCT 1333

Qy        393 TrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeuPhe 412
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1334 TGGAGACCTCAACGTGGTCTCCCAGAACGAGCCGTCTCAGTTCTCCGTGCTTGGCTCTTC 1393

Qy        413 GluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGlnThr 432
              ||||||||||||||||||||||||||||||||||||   ||||||||||||||||||||
Db       1394 GAACACTTTCTTCATCCATACCCTAAGGATTCGGACAAGCACATGCTAGCTAAGCAAACA 1453

Qy        433 GlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLys 452
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1454 GGACTCACTCGTAGCCAGGTGTCGAACTGGTTTATAAACGCGAGAGTTCGGTTATGGAAA 1513

Qy        453 ProMetValGluGluMetTyrLeuGluGluValLysAsnGlnGluGlnAsnSerThrAsn 472
              ||||||||||||||||||:::||||||:::|||   |||   :::|||
```

```
Db       1514 CCAATGGTGGAGGAGATGTACATGGAGGAAATGAAGGAGCAGGCAAAGAACATG------ 1567

Qy        473 ThrSerGlyAspAsnLysAsnLysGluThrAsnIleSerAlaProAsnGluGluLysHis 492

Db       1568 -----------GGATCCATGGAAAAGACTCCTTTGGATCAAAGCAACGAAGAT------ 1609

Qy        493 ProIleIleThrSerSerLeuLeuGlnAspGlyIleThrThrThrGlnAlaGluIleSer 512

Db       1610 ----------------------------------------------TCTGCTTCAAAGTCA 1624

Qy        513 ThrSerThrIleSerThrSerProThrAlaGlyAlaSerLeuHis------HisAlaHis 530

Db       1625 ACAAGTAACCAAGAAAAGAGCCCAATGGCGGACACTAATTACCATATGAATCCCAATCAC 1684

Qy        531 AsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrThrValAspHisIle 550

Db       1685 AACGGTGACCTAGAAGGCGTCACTGGAATGCAAGGATGCCCCAAGAGACTAAGAACCAGC 1744

Qy        551 GluAsnAsnAlaLysLysGlnArgAsnAspMetHisLysPheSerProSerSerIleLeu 570

Db       1745 GACGAGACAATGATGCAGCCAATAAAT------GCGGATTTCAGCTCCAACGAGAAGCTC 1798

Qy        571 SerSerValAspMetGluAlaLysAlaArgGluSerSerAsnLysGlyPheThrAsnPro 590

Db       1799 ACGATGAAGATTCTAGAAGAACGGCAAGGGATAAGATCAGATGGTGGCTAC------CCT 1852

Qy        591 LeuMet---------AlaAlaTyrAlaMetGlyAspPheGlyArgPheAsp---ProHis 606

Db       1853 TTCATGGGTAATTTCGGGCAATACCAAATGGATGAGATGTCAAGATTTGATGTAGTCTCA 1912

Qy        607 AspGlnGlnMetThrAlaAsnPheHis------GlyAsnAsnGlyValSerLeuThrLeu 624

Db       1913 GACCAGGAGCTCATGGCGCAAAGGTACTCAGGAAACAACAATGGCGTGTCCCTCACGTTA 1972

Qy        625 GlyLeuProProSerGluAsnLeuAlaMetProValSerGln---------GlnAsnTyr 641

Db       1973 GGTTTACCTCATTGTGATAGCTTGTCGTCCACGGACCATCAGGGTTTCATGCAGACCCAC 2032

Qy        642 LeuSerAsnAspLeuGlySerArgSerGluMetGly---------------------- 653

Db       2033 CATGGGATTCCTATAGGGAGAAGAGTGAAAATAGGAGAAACAGAGGAATATGGACCCGCC 2092

Qy        654 ----------------------------------------------------SerHis 655

Db       2093 ACCATCAATGGTGGTAGCTCGACCACAACCGCACATTCATCAGCGGCAGCTGCCGCGGCT 2152

Qy        656 TyrAsnArgMetGlyTyrGluAsnIleAspPheGlnSerGlyAsnLysArgProThr 675

Db       2153 TACAATGGGATGAACATACAGAAC----------------CAGAAGAGATATGTGGCT 2194

Qy        676 GlnLeuLeuProAspPheVal 682

Db       2195 CAGTTATTGCCCGACTTCGTT 2215
```

```
RESULT 7
US-10-374-780A-2671
; Sequence 2671, Application US/10374780A
; Publication No. US20040019927A1
; GENERAL INFORMATION:
;  APPLICANT: Sherman, Bradley K
;  APPLICANT:  Riechmann, Jose Luis
;  APPLICANT:  Jiang, Cai-Zhong
;  APPLICANT:  Heard, Jacqueline E
;  APPLICANT:  Haake, Volker
;  APPLICANT:  Creelman, Robert A
;  APPLICANT:  Ratcliffe, Oliver
;  APPLICANT:  Adam, Luc J
;  APPLICANT:  Reuber, T. Lynne
;  APPLICANT:  Keddie, James
;  APPLICANT:  Broun, Pierre E
;  APPLICANT:  Pilgrim, Marsha L
;  APPLICANT:  Dubell III, Arnold T
;  APPLICANT:  Pineda, Omaira
;  APPLICANT:  Yu, Guo-Liang
;  TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
;  FILE REFERENCE: MBI-0047 CIP
;  CURRENT APPLICATION NUMBER: US/10/374,780A
;  CURRENT FILING DATE:  2003-02-25
;  PRIOR APPLICATION NUMBER: 09/837,944
;  PRIOR FILING DATE: 2001-04-18
;  PRIOR APPLICATION NUMBER: 60/310,847
;  PRIOR FILING DATE: 2001-08-09
;  PRIOR APPLICATION NUMBER: 09/934,455
;  PRIOR FILING DATE: 2001-08-22
;  PRIOR APPLICATION NUMBER: 60/336,049
;  PRIOR FILING DATE: 2001-11-19
;  PRIOR APPLICATION NUMBER: 60/338,692
;  PRIOR FILING DATE: 2001-12-11
;  PRIOR APPLICATION NUMBER: 10/171,468
;  PRIOR FILING DATE: 2002-06-14
;  PRIOR APPLICATION NUMBER: 10/225,066
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: 10/225,067
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: 10/225,068
;  PRIOR FILING DATE: 2002-08-09
;  NUMBER OF SEQ ID NOS: 2906
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2671
;   LENGTH: 2385
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   OTHER INFORMATION: G1589
US-10-374-780A-2671

Alignment Scores:
Pred. No.:              4.81e-104       Length:         2385
Score:                  1148.50         Matches:        306
Percent Similarity:     51.67%          Conservative:   80
Best Local Similarity:  40.96%          Mismatches:     202
Query Match:            32.23%          Indels:         159
DB:                     17              Gaps:           28

US-10-624-201A-2 (1-688) x US-10-374-780A-2671 (1-2385)

Qy         22 HisGlyAsn-----------------SerAsnAsnAsnAsnIleGlnThrLeuTyrLeu 35

Db        194 CACGGAAACCCACCGGAGATCTCTGCCGGATCCGACGGTGGTCTTCAAACGTTGATCCTC 253

Qy         36 MetAsnProAsnAsnTyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGlnGlnGln 55

Db        254 ATGAATCCAACTACTTACGTTCAG---TACACCCAACAAGACAACGACTCGAACAACAAC 310

Qy         56 LeuLeuPheLeuAsnSerSerProAlaAlaSerAsnAlaLeuCysHisAlaAsnIle--- 74

Db        311 -----------AACAACAGCAACAATAGCAACAACAACAACACAAACACAAACACAAAC 358

Qy         75 ----------------------------GlnHisAlaProLeu------GlnGlnGln 82

Db        359 AACAACAACAGTAGTTTCGTTTTCCTCGATTCCCACGCGCCGCAGCCAAACGCGAGCCAG 418

Qy         83 HisPheValGlyValProLeu---------------ProAlaValSerLeuHisAspGln 97

Db        419 CAGTTCGTCGGAATACCACTCTCAGGTCACGAAGCTGCTTCCATTACAGCCGCCGACAAC 478

Qy         98 IleAsn---HisHisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnSerGlnGln 116

Db        479 ATCTCCGTACTTCACGGTTATCCTCCGCGCGTGCAGTACAGTCTCTACGGTAGCCACCAA 538

Qy        117 ValIleValProSerSerThrGlyValSerAlaThrSerCysGlyGlyIleThrThrAsp 136

Db        539 GTG---------------------------------------------------------- 541

Qy        137 LeuAlaSerGlnLeuAlaPheGlnArgProIleProThrProGlnHisArgGln------ 154
```

```
Db        542 ----------------------------------GATCCCACTCACCAGCAAGCCGCG 565

Qy        155 ------GlnGlnGlnGlnGlyGlyLeuSerLeuSerLeuSerProGlnLeuGlnGln 172

Db        566 TGTGAGACGCCACGCGCGCAGCAAGGCCTCTCTTTAACCCTCTCGTCTCAACAGCAGCAG 625

Qy        173 GlnIleSerPheAsnAsn--------------AsnIleSerSerSerSerProArgThr 187

Db        626 CAACAGCAACATCATCAACAACACCAGCCTATTCACGTCGGATTCGGGTCCGGACATGGA 685

Qy        188 AsnAsnValThrIleArgGlyThrLeuAspGlySerSer--------------SerAsn 202

Db        686 GAAGATATCCGGGTCGGGTCTGGCTCTACAGGATCGGGGGTAACAAACGGTATAGCTAAT 745

Qy        203 MetValLeuGlySerLysTyrLeuLysAlaAlaGlnGluLeuLeuAspGluValValAsn 222

Db        746 CTTGTT---AGCTCCAAGTACTTGAAGGCAGCACAAGAGCTTCTTGACGAAGTAGTCAAC 802

Qy        223 IleValGlyLysSerIleLysGlyAspAspGln------------LysLysAspAsnSer 238

Db        803 GCTGATTCCGATGACATGAACGCTAAATCCCAACTATTCTCATCGAAAAAGGGTAGTTGC 862

Qy        239 MetAsnLysGluSerMetProLeuAlaSer---AspValAsnThrAsnSerSerGlyGly 257

Db        863 GGAAATGATAAACCTGTCGGAGAATCATCGGCCGGCGCTGGAGGAGAAGGTTCCGGTGGC 922

Qy        258 GlyGluSerSerSerArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGlnArg 277

Db        923 GGAGCAGAAGCAGCCGGGAAACGT---------CCGGTGGAGCTAGGCACGGCAGAGAGA 973

Qy        278 GlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeuGluGluValGluGlnArg 297

Db        974 CAAGAAATACAGATGAAGAAAGCAAAACTTAGTAACATGCTTCATGAGGTGGAGCAGAGA 1033

Qy        298 TyrArgGlnTyrHisHisGlnMetGlnIleIleValLeuSerPheGluGlnValAlaGly 317

Db       1034 TATAGACAGTACCACCAGCAGATGCAGATGGTGATCTCTTCGTTCGAGCAAGCGGCAGGG 1093

Qy        318 IleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArg 337

Db       1094 ATAGGATCAGCGAAGTCATACACGTCGCTAGCATTGAAAACCATATCAAGACAGTTCCGT 1153

Qy        338 CysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGluGlu 357

Db       1154 TGCTTGAAAGAGGCGATCGCTGGTCAGATAAAAGCGGCCAACAAGAGTCTTGGGGAGGAA 1213

Qy        358 GluGlyLeuGly------GlyLysIleGluGlySerArgLeuLysPheValAspHisHis 375

Db       1214 GATTCAGTGTCTGGTGTTGGGAGGTTTGAGGGGTCGAGGCTCAAGTTCGTGGACCACCAC 1273

Qy        376 LeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGln---------ProAsnAla 392

Db       1274 TTGAGACAGCAAAGAGCTCTTCAACAACTGGGAATGATTCAACATCCTTCCAATAATGCT 1333

Qy        393 TrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeuPhe 412

Db       1334 TGGAGACCTCAACGTGGTCTCCCAGAACGAGCCGTCTCAGTTCTCCGTGCTTGGCTCTTC 1393

Qy        413 GluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGlnThr 432

Db       1394 GAACACTTTCTTCATCCATACCCTAAGGATTCGGACAAGCACATGCTAGCTAAGCAAACA 1453

Qy        433 GlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLys 452

Db       1454 GGACTCACTCGTAGCCAGGTGTCGAACTGGTTTATAAACGCGAGAGTTCGGTTATGGAAA 1513

Qy        453 ProMetValGluGluMetTyrLeuGluGluValLysAsnGlnGluGlnAsnSerThrAsn 472

Db       1514 CCAATGGTGGAGGAGATGTACATGGAGGAAATGAAGGAGCAGGCAAAGAACATG------ 1567

Qy        473 ThrSerGlyAspAsnLysAsnLysGluThrAsnIleSerAlaProAsnGluGluLysHis 492

Db       1568 ------------GGATCCATGGAAAAGACTCCTTTGGATCAAAGCAACGAAGAT------ 1609
```

```
Qy        493 ProIleIleThrSerSerLeuLeuGlnAspGlyIleThrThrThrGlnAlaGluIleSer 512

Db       1610 ----------------------------------------------TCTGCTTCAAAGTCA 1624

Qy        513 ThrSerThrIleSerThrSerProThrAlaGlyAlaSerLeuHis------HisAlaHis 530

Db       1625 ACAAGTAACCAAGAAAAGAGCCCAATGGCGGACACTAATTACCATATGAATCCCAATCAC 1684

Qy        531 AsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrThrValAspHisIle 550

Db       1685 AACGGTGACCTAGAAGGCGTCACTGGAATGCAAGGATGCCCCAAGAGACTAAGAACCAGC 1744

Qy        551 GluAsnAsnAlaLysLysGlnArgAsnAspMetHisLysPheSerProSerSerIleLeu 570

Db       1745 GACGAGACAATGATGCAGCCAATAAAT------GCGGATTTCAGCTCCAACGAGAAGCTC 1798

Qy        571 SerSerValAspMetGluAlaLysAlaArgGluSerSerAsnLysGlyPheThrAsnPro 590

Db       1799 ACGATGAAGATTCTAGAAGAACGGCAAGGGATAAGATCAGATGGTGGCTAC------CCT 1852

Qy        591 LeuMet---------AlaAlaTyrAlaMetGlyAspPheGlyArgPheAsp---ProHis 606

Db       1853 TTCATGGGTAATTTCGGGCAATACCAAATGGATGAGATGTCAAGATTTGATGTAGTCTCA 1912

Qy        607 AspGlnGlnMetThrAlaAsnPheHis------GlyAsnAsnGlyValSerLeuThrLeu 624

Db       1913 GACCAGGAGCTCATGGCGCAAAGGTACTCAGGAAACAACAATGGCGTGTCCCTCACGTTA 1972

Qy        625 GlyLeuProProSerGluAsnLeuAlaMetProValSerGln---------GlnAsnTyr 641

Db       1973 GGTTTACCTCATTGTGATAGCTTGTCGTCCACGGACCATCAGGGTTTCATGCAGACCCAC 2032

Qy        642 LeuSerAsnAspLeuGlySerArgSerGluMetGly----------------------- 653

Db       2033 CATGGGATTCCTATAGGGAGAAGAGTGAAAATAGGAGAAACAGAGGAATATGGACCCGCC 2092

Qy        654 ----------------------------------------------------SerHis 655

Db       2093 ACCATCAATGGTGGTAGCTCGACCACAACCGCACATTCATCAGCGGCAGCTGCCGCGGCT 2152

Qy        656 TyrAsnArgMetGlyTyrGluAsnIleAspPheGlnSerGlyAsnLysArgPheProThr 675

Db       2153 TACAATGGGATGAACATACAGAAC-----------------CAGAAGAGATATGTGGCT 2194

Qy        676 GlnLeuLeuProAspPheVal 682

Db       2195 CAGTTATTGCCCGACTTCGTT 2215
```

RESULT 8
US-10-225-066A-335
; Sequence 335, Application US/10225066A
; Publication No. US20050160493A9
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: RATCLIFFE, Oliver
; APPLICANT: RIECHMANN, Jose Luis
; APPLICANT: ADAM, Luc J
; APPLICANT: DUBELL, Arnold T
; APPLICANT: HEARD, Jacqueline E
; APPLICANT: PILGRIM, Marsha L
; APPLICANT: JIANG, Cai-Zhong
; APPLICANT: REUBER, T. Lynne
; APPLICANT: CREELMAN, Robert A
; APPLICANT: PINEDA, Omaira
; APPLICANT: YU, Guo-Liang
; APPLICANT: BROUN, Pierre E
; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: MBI0036-2 US
; CURRENT APPLICATION NUMBER: US/10/225,066A
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 09/837,444
; PRIOR FILING DATE: 2001-04-18

```
;  PRIOR APPLICATION NUMBER: 60/310,847
;  PRIOR FILING DATE: 2001-08-09
;  PRIOR APPLICATION NUMBER: 60/336,049
;  PRIOR FILING DATE: 2001-12-05
;  PRIOR APPLICATION NUMBER: 60/338,692
;  PRIOR FILING DATE: 2001-12-11
;  PRIOR APPLICATION NUMBER: 10/171,468
;  PRIOR FILING DATE: 2002-06-14
;  NUMBER OF SEQ ID NOS: 1122
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 335
;   LENGTH: 2385
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-10-225-066A-335

Alignment Scores:
Pred. No.:                4.81e-104      Length:          2385
Score:                    1148.50        Matches:         306
Percent Similarity:       51.67%         Conservative:    80
Best Local Similarity:    40.96%         Mismatches:      202
Query Match:              32.23%         Indels:          159
DB:                       22             Gaps:            28

US-10-624-201A-2 (1-688) x US-10-225-066A-335 (1-2385)

Qy         22 HisGlyAsn-----------------SerAsnAsnAsnAsnIleGlnThrLeuTyrLeu 35
              |||||||||                :::::     :::|||||||||    |||
Db        194 CACGGAAACCCACCGGAGATCTCTGCCGGATCCGACGGTGGTCTTCAAACGTTGATCCTC 253

Qy         36 MetAsnProAsnAsnTyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGlnGlnGln 55
              |||||||||      |||:::|||   ||||||      |||
Db        254 ATGAATCCAACTACTTACGTTCAG---TACACCCAACAAGACAACGACTCGAACAACAAC 310

Qy         56 LeuLeuPheLeuAsnSerSerProAlaAlaSerAsnAlaLeuCysHisAlaAsnIle--- 74
                  |||:::|||      ::::::|||         :::   |||
Db        311 -----------AACAACAGCAACAATAGCAACAACAACAACACAAACACAAACACAAAC 358

Qy         75 ------------------------------GlnHisAlaProLeu------GlnGlnGln 82
                                            |||||||||                |||
Db        359 AACAACAACAGTAGTTTCGTTTTCCTCGATTCCCACGCGCCGCAGCCAAACGCGAGCCAG 418

Qy         83 HisPheValGlyValProLeu---------------ProAlaValSerLeuHisAspGln 97
                 |||||||||:::||||||              :::::::::      |||
Db        419 CAGTTCGTCGGAATACCACTCTCAGGTCACGAAGCTGCTTCCATTACAGCCGCCGACAAC 478

Qy         98 IleAsn---HisHisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnSerGlnGln 116
              |||:::     ||||||          |||:::      :::          |||   |||
Db        479 ATCTCCGTACTTCACGGTTATCCTCCGCGCGTGCAGTACAGTCTCTACGGTAGCCACCAA 538

Qy        117 ValIleValProSerSerThrGlyValSerAlaThrSerCysGlyGlyIleThrThrAsp 136
              |||
Db        539 GTG------------------------------------------------------- 541

Qy        137 LeuAlaSerGlnLeuAlaPheGlnArgProIleProThrProGlnHisArgGln------ 154
                                                |||   |||:::|||
Db        542 -----------------------------------GATCCCACTCACCAGCAAGCCGCG 565

Qy        155 ------GlnGlnGlnGlnGlnGlyGlyLeuSerLeuSerLeuSerProGlnLeuGlnGln 172
                    :::   |||   ||||||||||||:::||||||   |||   ||||||
Db        566 TGTGAGACGCCACGCGCGCAGCAAGGCCTCTCTTTAACCCTCTCGTCTCAACAGCAGCAG 625

Qy        173 GlnIleSerPheAsnAsn---------------AsnIleSerSerSerSerProArgThr 187
              |||          :::                ::::::          |||
Db        626 CAACAGCAACATCATCAACAACACCAGCCTATTCACGTCGGATTCGGGTCCGGACATGGA 685

Qy        188 AsnAsnValThrIleArgGlyThrLeuAspGlySerSer---------------SerAsn 202
                 ::::::   :::                |||||                 :::|||
Db        686 GAAGATATCCGGGTCGGGTCTGGCTCTACAGGATCGGGGGTAACAAACGGTATAGCTAAT 745

Qy        203 MetValLeuGlySerLysTyrLeuLysAlaAlaGlnGluLeuLeuAspGluValValAsn 222
              :::|||      ||||||||||||||||||||||||||||||||||||||||||||||
Db        746 CTTGTT---AGCTCCAAGTACTTGAAGGCAGCACAAGAGCTTCTTGACGAAGTAGTCAAC 802

Qy        223 IleValGlyLysSerIleLysGlyAspAspGln-----------LysLysAspAsnSer 238
                          :::             |||           ||||||   :::
Db        803 GCTGATTCCGATGACATGAACGCTAAATCCCAACTATTCTCATCGAAAAAGGGTAGTTGC 862

Qy        239 MetAsnLysGluSerMetProLeuAlaSer---AspValAsnThrAsnSerSerGlyGly 257
                 |||   :::   :::     :::|||                        |||||||||
Db        863 GGAAATGATAAACCTGTCGGAGAATCATCGGCCGGCGCTGGAGGAGAAGGTTCCGGTGGC 922

Qy        258 GlyGluSerSerSerArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGlnArg 277
                 |||      ::::::   ::::::           |||||||||   ||||||:::|||
Db        923 GGAGCAGAAGCAGCCGGGAAACGT---------CCGGTGGAGCTAGGCACGGCAGAGAGA 973

Qy        278 GlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeuGluGluValGluGlnArg 297
              ||||||:::||||||||||||||||||||||      ||||||   ||||||||||||||||
Db        974 CAAGAAATACAGATGAAGAAAGCAAAACTTAGTAACATGCTTCATGAGGTGGAGCAGAGA 1033

Qy        298 TyrArgGlnTyrHisHisGlnMetGlnIleIleValLeuSerPheGluGlnValAlaGly 317
              ||||||||||||||    |||||||||::::::::    ||||||||||    ||||||
Db       1034 TATAGACAGTACCACCAGCAGATGCAGATGGTGATCTCTTCGTTCGAGCAAGCGGCAGGG 1093

Qy        318 IleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArg 337
              ||||||||||||||||||||||||    |||||||||      ||||||:::|||||||||
Db       1094 ATAGGATCAGCGAAGTCATACACGTCGCTAGCATTGAAAACCATATCAAGACAGTTCCGT 1153

Qy        338 CysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGluGlu 357
              |||||||||:::|||||||||   |||:::||||||   :::|||||||||||||||||||
Db       1154 TGCTTGAAAGAGGCGATCGCTGGTCAGATAAAAGCGGCCAACAAGAGTCTTGGGGAGGAA 1213

Qy        358 GluGlyLeuGly------GlyLysIleGluGlySerArgLeuLysPheValAspHisHis 375
              :::   :::         |||:::   ||||||||||||||||||||||||||||||||
Db       1214 GATTCAGTGTCTGGTGTTGGGAGGTTTGAGGGGTCGAGGCTCAAGTTCGTGGACCACCAC 1273

Qy        376 LeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGln---------ProAsnAla 392
              ||||||||||||||||||||||||||:::||||||:::|||          ||||||
Db       1274 TTGAGACAGCAAAGAGCTCTTCAACAACTGGGAATGATTCAACATCCTTCCAATAATGCT 1333

Qy        393 TrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeuPhe 412
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1334 TGGAGACCTCAACGTGGTCTCCCAGAACGAGCCGTCTCAGTTCTCCGTGCTTGGCTCTTC 1393

Qy        413 GluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGlnThr 432
              ||||||||||||||||||||||||||||||||||||   ||||||||||||||||||
Db       1394 GAACACTTTCTTCATCCATACCCTAAGGATTCGGACAAGCACATGCTAGCTAAGCAAACA 1453

Qy        433 GlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLys 452
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1454 GGACTCACTCGTAGCCAGGTGTCGAACTGGTTTATAAACGCGAGAGTTCGGTTATGGAAA 1513

Qy        453 ProMetValGluGluMetTyrLeuGluGluValLysAsnGlnGluGlnAsnSerThrAsn 472
              ||||||||||||||||||:::||||||:::|||   |||   :::|||
Db       1514 CCAATGGTGGAGGAGATGTACATGGAGGAAATGAAGGAGCAGGCAAAGAACATG------ 1567

Qy        473 ThrSerGlyAspAsnLysAsnLysGluThrAsnIleSerAlaProAsnGluGluLysHis 492
                                  ::::::|||   :::   ||||||:::
Db       1568 -----------GGATCCATGGAAAAGACTCCTTTGGATCAAAGCAACGAAGAT------ 1609

Qy        493 ProIleIleThrSerSerLeuLeuGlnAspGlyIleThrThrThrGlnAlaGluIleSer 512
                                                        |||       |||
Db       1610 ------------------------------------------TCTGCTTCAAAGTCA 1624

Qy        513 ThrSerThrIleSerThrSerProThrAlaGlyAlaSerLeuHis------HisAlaHis 530
              ||||||              ||||||   |||      :::   |||           |||
Db       1625 ACAAGTAACCAAGAAAAGAGCCCAATGGCGGACACTAATTACCATATGAATCCCAATCAC 1684

Qy        531 AsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrThrValAspHisIle 550
              |||          |||          |||                    :::
Db       1685 AACGGTGACCTAGAAGGCGTCACTGGAATGCAAGGATGCCCCAAGAGACTAAGAACCAGC 1744

Qy        551 GluAsnAsnAlaLysLysGlnArgAsnAspMetHisLysPheSerProSerSerIleLeu 570
              :::          :::      |||          ||||||   :::      |||
Db       1745 GACGAGACAATGATGCAGCCAATAAAT------GCGGATTTCAGCTCCAACGAGAAGCTC 1798
```

```
Qy        571 SerSerValAspMetGluAlaLysAlaArgGluSerSerAsnLysGlyPheThrAsnPro 590
              :::         :::|||   :::             |||:::   |||:::      |||
Db       1799 ACGATGAAGATTCTAGAAGAACGGCAAGGGATAAGATCAGATGGTGGCTAC------CCT 1852

Qy        591 LeuMet---------AlaAlaTyrAlaMetGlyAspPheGlyArgPheAsp---ProHis 606
                 |||             |||   |||   :::     ||||||||          
Db       1853 TTCATGGGTAATTTCGGGCAATACCAAATGGATGAGATGTCAAGATTTGATGTAGTCTCA 1912

Qy        607 AspGlnGlnMetThrAlaAsnPheHis------GlyAsnAsnGlyValSerLeuThrLeu 624
              ||||||:::::   |||      :::         ||||||||||||||||||||||||
Db       1913 GACCAGGAGCTCATGGCGCAAAGGTACTCAGGAAACAACAATGGCGTGTCCCTCACGTTA 1972

Qy        625 GlyLeuProProSerGluAsnLeuAlaMetProValSerGln---------GlnAsnTyr 641
              |||||||||      ::::::|||:::             |||         |||   :::
Db       1973 GGTTTACCTCATTGTGATAGCTTGTCGTCCACGGACCATCAGGGTTTCATGCAGACCCAC 2032

Qy        642 LeuSerAsnAspLeuGlySerArgSerGluMetGly----------------------- 653
                        :::|||   |||   ::::::|||
Db       2033 CATGGGATTCCTATAGGGAGAAGAGTGAAAATAGGAGAAACAGAGGAATATGGACCCGCC 2092

Qy        654 ------------------------------------------------------SerHis 655
                                                                     :::
Db       2093 ACCATCAATGGTGGTAGCTCGACCACAACCGCACATTCATCAGCGGCAGCTGCCGCGGCT 2152

Qy        656 TyrAsnArgMetGlyTyrGluAsnIleAspPheGlnSerGlyAsnLysArgProProThr 675
              ||||||   |||      :::|||                     |||||||:::
Db       2153 TACAATGGGATGAACATACAGAAC-----------------CAGAAGAGATATGTGGCT 2194

Qy        676 GlnLeuLeuProAspPheVal 682
              |||||||||||||||||||
Db       2195 CAGTTATTGCCCGACTTCGTT 2215

RESULT 9
US-10-424-599-133558
; Sequence 133558, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 133558
;   LENGTH: 1822
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_91608C.1
US-10-424-599-133558

Alignment Scores:
Pred. No.:              2.58e-97      Length:        1822
Score:                  1079.50       Matches:       248
Percent Similarity:     64.19%        Conservative:  55
Best Local Similarity:  52.54%        Mismatches:    112
Query Match:            30.30%        Indels:        57
DB:                     18            Gaps:          16

US-10-624-201A-2 (1-688) x US-10-424-599-133558 (1-1822)

Qy        253 AsnSerSerGlyGlyGlyGluSerSerSerArgGlnLys---AsnGluValAlaValGlu 271
              :::|||   |||||||||||:::::::::          :::       ||||||
Db         12 TCGAGCTGTGGAGGAGGGGAGAATAATAATAATAATGATGGAGGCAAACAAGGAGTTGAA 71

Qy        272 LeuThrThrAlaGlnArgGlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeu 291
              |||:::||||||||||||||||||||||||||||:::|||||:::   ||||||
Db         72 CTCAGCACAGCACAGAGACAAGAGCTTCAGATGAAGAAGTCAAAACTTGTGACCATGCTA 131

Qy        292 GluGluValGluGlnArgTyrArgGlnTyrHisHisGlnMetGlnIleIleValLeuSer 311
              :::|||||||||||||||   |||||||||||||||||||||:::::::::   |||
Db        132 GATGAGGTGGAGCAAAGGGACCGACAGTATCACCACCAAATGCAAGTTGTGATAACATCA 191

Qy        312 PheGluGlnValAlaGlyIleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAla 331
              |||||||||   ||||||:::|||:::|||||||||||||||   |||||||||
Db        192 TTTGAGCAAGCAGCGGGTGTTGGAGCTGCAAAGTCTTACACAGCCCTTGCGTTAAAGACA 251

Qy        332 IleSerLysGlnPheArgCysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSer 351
              |||||||||||||||||||||||||||||||||||:::   |||:::|||   ||||||
Db        252 ATCTCAAAGCAATTTCGGTGTCTCAAAGATGCAATCTCTTCACAAATCAAGACAACTAGC 311

Qy        352 LysSerLeuGlyGluGluGluGlyLeuGlyGlyLysIleGluGlySerArgLeuLysPhe 371
              |||:::|||||||||:::      ||||||   |||:::|||||||||||||||||||||
Db        312 AAGACCTTAGGTGAAGATAATTGCCTGGGAGTTAAGGTAGAGGGTTCGAGGCTAAGGTAT 371

Qy        372 ValAspHisHisLeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGlnProAsn 391
              |||||||||      ||||||||||||||||||   |||:::||||||:::|||   |||
Db        372 GTTGATCATCAACAGAGACAACAACGTGCATTA---CAGCTTGGAATGATCCAACACAAC 428

Qy        392 AlaTrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeu 411
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        429 GCTTGGAGGCCCCAAAGAGGCTTGCCTGAACGTGCTGTCTCTGTCCTTCGTGCTTGGCTT 488

Qy        412 PheGluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGln 431
              |||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||||
Db        489 TTTGAACATTTCTTGCACCCCTATCCTAAGGACTCCGATAAGGTTATGCTTGCAAAACAA 548

Qy        432 ThrGlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrp 451
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        549 ACTGGGCTTACTCGGAGCCAGGTATCTAACTGGTTTATAAATGCCCGAGTTCGGCTATGG 608

Qy        452 LysProMetValGluGluMetTyrLeuGluGluValLys-----------AsnGlnGlu 467
              |||||||||||||||||||||||||||||||||||||||          |||
Db        609 AAGCCAATGGTAGAAGAAATGTACTTGGAAGAGGTAAAACAAGAGCCAAACAATAGCTCC 668

Qy        468 GlnAsnSerThrAsnThrSerGlyAspAsnLysAsnLysGluThrAsnIleSerAlaPro 487
              |||::::::   ||||||   |||   |||::::::   :::         |||
Db        669 CAGGATAACAACAACACAAAAGGATCAAATGAATCAAGCAAGGAGTTATGGTCAGAAGCT 728

Qy        488 AsnGluGluLysHisProIleIleThrSerSerLeuLeuGlnAspGlyIleThrThrThr 507
              |||                                     |||      |||      :::
Db        729 AATGCTGCTGCACAAGAATCAGGTGCAATGAGATTTGATCAGATTAACATTCTCCAATCA 788

Qy        508 GlnAlaGluIleSerThrSerThrIleSerThrSerProThrAlaGlyAlaSerLeuHis 527
              :::||||||         :::      :::||||||||||||
Db        789 AAGGCAGAAAGCTTCAAGAACAACCAAACCACTTCCCCAACA--------------GAG 833

Qy        528 HisAlaHisAsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrThrVal 547
                 :::::::::   |||            |||::::::
Db        834 ATCTCCAATTCCAACTCACTCCAATCTGGTTTCCATCTT----------------GCT 875

Qy        548 AspHisIleGluAsnAsnAlaLysLysGlnArgAsnAspMetHisLysPheSerProSer 567
              |||            |||      |||      ::::::||||||:::   ||||||
Db        876 GACATGCAAAGTCCAAACAAGCCAAAAAGCACTTCTGAGATGCATCAGAATTCTCCAGGT 935

Qy        568 SerIleLeuSerSerValAspMetGluAlaLys--------------------------- 578
              |||||||||   ||||||||||||||   |||
Db        936 AGCATTCTT---TCAGTGGACATGGAAATGAAGCCTCATCATCATGGAGAAACAAATAAT 992

Qy        579 ---AlaArgGluSerSerAsnLysGlyPheThrAsnProLeu------------------ 591
                    ||||||      |||         |||
Db        993 ATCACTAGAGAAGGCCAGAACAACAACACAACAACAAAGTTTGGCATTGAAAGCCATGGT 1052

Qy        592 ------MetAlaAlaTyrAla---MetGlyAspPheGlyArgPheAspProHis---Asp 607
                   |||:::      |||   |||   |||||||||      |||
Db       1053 GGGGGTTTTGGAGCATTTCCCAATATGGAAGACATTGGAAGGTTTCATCATCATGTCACA 1112

Qy        608 GlnGlnMetThrAlaAsnPheHisGlyAsnAsnGlyValSerLeuThrLeuGlyLeuPro 627
              :::|||:::         |||||||||   ||||||||||||||||||||||||||||
Db       1113 GAGCAACTTGCTCCTAGGTTCCATGGA---AATGGTGTTTCTCTCACTCTTGGACTTCCA 1169

Qy        628 ProSer------GluAsnLeuAlaMetPro---ValSerGlnGlnAsnTyrLeuSerAsn 644
```

```
Db       1170 CACAGTACTGAGAATAACCTTTCCCTCTCAGGAACAACCCAACATGGATTCCTCTCACAG 1229

Qy        645 Asp------LeuGlySerArgSerGluMetGly---------------------------- 653

Db       1230 AACATGCACTTGGGAATGAGAAACAGTGAAAATGATGAGTTTTGTGGTGCAATCAACACT 1289

Qy        654 ---------SerHisTyrAsnArgMetGlyTyrGluAsnIleAspPheGlnSerGlyAsn 670

Db       1290 ACCCCACCTTCTTCTCACTCAGGTATCAGTTATGAGAGCATTGACATTCAAAAC---AGA 1346

Qy        671 LysArgProThrGlnLeuLeuProAspPheVal 682

Db       1347 AAGAGGTTTGCTGCGCAGTTGTTGCGGGATTTCGTG 1382

RESULT 10
US-10-425-114-7804
; Sequence 7804, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 7804
;   LENGTH: 1606
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: 700684286_FLI
US-10-425-114-7804

Alignment Scores:
Pred. No.:              3.02e-97      Length:          1606
Score:                  1078.00       Matches:         259
Percent Similarity:     64.10%        Conservative:    57
Best Local Similarity:  52.54%        Mismatches:      102
Query Match:            30.26%        Indels:          76
DB:                     18            Gaps:            18

US-10-624-201A-2 (1-688) x US-10-425-114-7804 (1-1606)

Qy        253 AsnSerSerGlyGlyGlyGluSerSerSerArgGlnLys---AsnGluValAlaValGlu 271

Db          3 TCGAGCTGTGGAGGAGGGGAGAATAATAATAATGATGGAGGCAAACAAGGAGTTGAA 62

Qy        272 LeuThrThrAlaGlnArgGlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeu 291

Db         63 CTCAGCACAGCACAGAGACAAGAGCTTCAGATGAAGAAGTCAAAACTTGTGACCATGCTA 122

Qy        292 GluGluValGluGlnArgTyrArgGlnTyrHisHisGlnMetGlnIleIleValLeuSer 311

Db        123 GATGAGGTGGAGCAAAGGTACCGACAGTATCACCACCAAATGCAAGTTGTGATAACATCA 182

Qy        312 PheGluGlnValAlaGlyIleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAla 331

Db        183 TTTGAGCAAGCAGCGGGTGTTGGAGCTGCAAAGTCTTACACAGCCCTTGCGTTAAAGACA 242

Qy        332 IleSerLysGlnPheArgCysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSer 351

Db        243 ATCTCAAAGCAATTTCGGTGTCTCAAAGATGCAATCTCTTCACAAATAAAGACCACGAGC 302

Qy        352 LysSerLeuGlyGluGluGluGlyLeuGlyGlyLysIleGluGlySerArgLeuLysPhe 371

Db        303 AAAACCTTAGGTGAAGATGATTGCTTGGGAGTTAAGGTAGAAGGTTCGAGGCTTAGGTAT 362

Qy        372 ValAspHisHisLeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGlnProAsn 391

Db        363 GTTGATCATCAACTGAGACAACAACGTGCATTACAACAGCTTGGAATGATCCAACACAAT 422

Qy        392 AlaTrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeu 411

Db        423 GCTTGGAGGCCCCAAAGAGGCTTGCCTGAACGTGCTGTTTCTGTTCTTCGAGCTTGGCTT 482

Qy        412 PheGluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGln 431

Db        483 TTTGAACATTTCTTGCACCCTTATCCTAAGGACTCCGATAAGGTTATGCTTGCAAAACAA 542

Qy        432 ThrGlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrp 451

Db        543 ACTGGACTTACTCGGAGCCAGGTGTCTAACTGGTTTATAAATGCCCGAGTTCGGCTATGG 602

Qy        452 LysProMetValGluGluMetTyrLeuGluGluValLysAsnGlnGluGlnAsnSerThr 471

Db        603 AAGCCAATGGTAGAAGAAATGTACTTGGAAGAGGTAAAACAAGAACCAAACAATAGCTCT 662

Qy        472 ------AsnThrSer-GlyAspAsnLysAsnLysGluThrAsnIleSerAlaProAsnGl 489

Db        663 CAGGATAACACCACAAAAAGATCAAAAGAATCAAGCAAGGAGTTATGGT----------- 711

Qy        489 uGluLysHisProIleIleThrSerSerLeuLeuGln---AspGlyIleThrThrThr-- 507

Db        712 -CAGAAGCTAATGCTACTGCACAGGAATCTGGTGCAATGAGACTTGATCACATCAACATT 770

Qy        508 ----------GlnAlaGluIleSerThrSerThrIleSerThrSerPro---ThrAlaGl 523

Db        771 CTCCAATCAAAAGCAGAAAGCTTCAACAACAACAACAACAACCAAACCACTTCCCCAACA 830

Qy        523 yAlaSerLeuHisHisAlaHisAsnPheSerPheLeuGlySerPheAsnMetAspAsnTh 543

Db        831 GAGATCTCTAAT-AATTCCCATAATTCACTCCAATCAGCTGGTTTCCATCTTGCTGACAT 889

Qy        543 rThrThrThrValAspHisIleGluAsnAsnAlaLysLysGlnArgAsnAspMetHisLy 563

Db        890 GCAAAGCCCAAACAAGCCAATTAGAAGCACTTCTGAAATGCAGAAC-------------- 935

Qy        563 sPheSerProSerSerIleLeuSerSerValAspMetGluAlaLysAlaArg-------- 580

Db        936 ----TCTCCAGGTAGCATTCTT---TCAGTGGACATGGAAATGAAGCCTCGTCATCATGG 988

Qy        581 -GluSerSerAsn-----------------LysGlyPheThrAsnProLeuMet----- 592

Db        989 AGAGACAAATAATAATACTATTATTACTAGAGAAGGCAACAACAACACAACAAAGTTTGG 1048

Qy        593 -----------------------------------AlaAlaTyrAlaMetGlyAspPh 600

Db       1049 CATTGAAAGCCATGGTGGTGGTGGTGGTTTTGGAGCATTTCCCAATATGGAAGACAT 1108

Qy        600 eGlyArgPheAspProHis--------------AspGlnGlnMetThrAla---AsnPh 614

Db       1109 TGGAAGGTTCCATCATCATCATCATCATGTCACTGATCAGCAATCACTGGCTCCTAGGTT 1168

Qy        614 eHisGlyAsnAsnGlyValSerLeuThrLeuGlyLeuProProSer---GluAsn----- 631

Db       1169 TCATGGA---AATGGTGTTTCTCTCACTCTTGGACTTCCACACAGCACTGAGAATAATAA 1225

Qy        632 ----------LeuAlaMetProValSerGlnGlnAsnTyrLeuSerAsnAsp------Le 646

Db       1226 TAACCTTTCCCTCTCAGGAACTACTCATCAACATGGATTCCTCTCACAGAACATGCACTT 1285

Qy        646 uGlySerArgSerGluMetGlySerHisTyrAsnArgMet------------------- 659

Db       1286 GGGAATGAGATCAACAACCAATAATGCGAATAATGAGTTTTGTGGTGCCATCAACACTAC 1345

Qy        660 --------------------------GlyTyrGluAsnIleAspPheGlnSerGlyAs 670

Db       1346 TCCACCTTCTAATTCTCACTCAGGCACCAGTTATGAGAGCATTGACATTATTCAAAACAG 1405
```

```
Qy        670 nLysArgPheProThrGlnLeuLeuProAspPheVal 682
Db       1406 AAAGAGATTTGCTGCGCAGTTGTTGCGGGATTTCGTG 1442

RESULT 11
US-10-437-963-60154
; Sequence 60154, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  APPLICANT:  Wu, Wei
;  APPLICANT:  Boukharov, Andrey A.
;  APPLICANT:  Barbazuk, Brad
;  APPLICANT:  Li, Ping
;  TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53221)B
;  CURRENT APPLICATION NUMBER: US/10/437,963
;  CURRENT FILING DATE:  2003-05-14
;  NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 60154
;   LENGTH: 3029
;   TYPE: DNA
;   ORGANISM: Oryza sativa
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT4530_6170C.1
US-10-437-963-60154

Alignment Scores:
Pred. No.:              4.41e-94      Length:        3029
Score:                  1050.50       Matches:       294
Percent Similarity:     50.98%        Conservative:  95
Best Local Similarity:  38.53%        Mismatches:    204
Query Match:            29.48%        Indels:        171
DB:                     19            Gaps:          27

US-10-624-201A-2 (1-688) x US-10-437-963-60154 (1-3029)

Qy          2 TyrTyrGlnGlyThrSerAspAsnThrAsnIleGlnAlaAspHisGlnGlnArgHisAsn 21
Db       1002 TACTACCACGGCGGCGCC---GGCACGGACATCCAG---------------------- 1034

Qy         22 HisGlyAsnSerAsnAsnAsnAsnIleGlnThrLeuTyrLeuMetAsnProAsnAsnTyr 41
Db       1035 ---------TCCGGCACCGACGGCCTGCAGACGCTGTACCTCATGAACCCGAGC------ 1079

Qy         42 MetGlnGlyTyr---------------ThrThrSerAspThrGlnGlnGlnGlnGlnLeu 56
Db       1080 TACGCCGGCTACGGCGACGCCGCCGCCGCCGCCGCCGCCGGGGCGGCGGCCAACATG 1139

Qy         57 LeuPheLeuAsnSer-----------------------------------------Ser 62
Db       1140 ATGCTCCTGAACTCGGCGGTGACCTCCATGACGCCGGTGTCGTTCGGCCACCAGCCGTCG 1199

Qy         63 ProAlaAlaSerAsnAlaLeuCysHisAlaAsnIleGlnHisAlaProLeuGlnGlnGln 82
Db       1200 CCGTCGTCGTCGTCGGCG-----------------------------------GCGCAG 1223

Qy         83 HisPheValGlyValProLeu------ProAlaValSerLeuHisAspGlnIleAsnHis 100
Db       1224 CACTTCGTCGGCATCCCTCTCCAGGCGCCGCCGGCTTCCGGGTACAAC------------ 1271

Qy        101 HisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnSerGlnGlnValIleValPro 120
Db       1272 -----------------CTGTGG---------------------------------- 1277

Qy        121 SerSerThrGlyValSerAlaThrSerCysGlyGlyIleThrThrAspLeuAlaSerGln 140
Db       1278 ------ACCCCGGCGGCGGCCACCGGCGCGGGCGACATGTCGCCG-------------- 1316

Qy        141 LeuAlaPheGlnArgProIleProThrProGlnHisArgGlnGlnGlnGlnGlnGlnGly 160
Db       1317 --------------------CCGACGCCGCAGCACCAGCATCAGCAAGCTCACGGCGGT 1355

Qy        161 GlyLeuSer--------------LeuSerLeuSerProGlnLeuGlnGlnGlnIleSer 175
Db       1356 GGCGCGGCGGGTGTCAGTGCTGTCCTCAGCCTGTCGTCCCGC----------------- 1397

Qy        176 PheAsnAsnAsnIleSerSerSerSerProArgThrAsnAsnValThrIleArgGlyThr 195
Db       1398 --------------GAGGCGGCGCCGCCGGTGACGGTGGCGGCCGTGGTGGCCGCCGGC 1442

Qy        196 LeuAspGly---------------------------------SerSerSerAsnMetVal 204
Db       1443 GACGAGGGGAAGTACCTGCAGGCGGTGGCGCAGGGCGCGGCGTCGCACGGGCAGATGGTG 1502

Qy        205 LeuGlySerLysTyrLeuLysAlaAlaGlnGluLeuLeuAspGluValValAsnIleVal 224
Db       1503 ATGAGCTCCAAGTACCTCAAGGCGGCGCAGGAGCTGCTCGACGAGGTGGTGAGCGTC--- 1559

Qy        225 GlyLysSerIleLysGlyAspAspGlnLysLysAspAsnSerMetAsnLysGlu------ 242
Db       1560 ---------AGCAAGGGCGTGGACGACGTCAAGGCCGCCGCGGCGGCGAAGAGCCCGGCC 1610

Qy        243 SerMetProLeuAlaSerAspValAsnThrAsnSerSerGlyGlyGlyGluSerSerSer 262
Db       1611 TCGGTGAAGAAGAAGGAGGACTCGGAGGGCGTGTCCGGCGGCGGCACGGAGGATGGCGGC 1670

Qy        263 ArgGlnLysAsnGluValAla---------ValGluLeuThrThrAlaGlnArgGlnGlu 279
Db       1671 GGCGCCAAGAGCGGCGGCGCGCCGCCGCCGCCGGAGATGTCGACGGCGGAGCGGCAGGAG 1730

Qy        280 LeuGlnMetLysLysAlaLysLeuLeuAlaMetLeuGluGluValGluGlnArgTyrArg 299
Db       1731 CTGCAGATGAAGAAAGGAAAGCTGATTAACATGCTTGATGAGGTGGAGCAGCGGTACAGG 1790

Qy        300 GlnTyrHisHisGlnMetGlnIleIleValLeuSerPheGluGlnValAlaGlyIleGly 319
Db       1791 CAGTACCACCAGCAGATGCAGGTGGTGGTGGCGTCGTTCGAGGCGGTGGCGGGGGGCGGG 1850

Qy        320 SerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArgCysLeu 339
Db       1851 TCGGCGAGGACGTACACGGCGCTGGCGCTGCGGACCATCTCGCGGCAGTTCCGGTGCCTG 1910

Qy        340 LysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGlu--------- 356
Db       1911 CGGGACGCGATCGCGGGGCAGGTGAGGGCGGCGAGCCGGGCGCTGGGGGAGGCCGTCGAC 1970

Qy        357 GluGluGlyLeuGlyGlyLysIleGluGlySerArgLeuLysPheValAspHisHisLeu 376
Db       1971 GCCGACGGCGGATGCGGCCGCACGGTGGGGTCCAGGCTCCGGTACATCGACCACCAGCTC 2030

Qy        377 ArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGlnProAsnAlaTrpArgProGln 396
Db       2031 CGGCAGCAGCGCGCGCTGCAGCAGCTGGGCATGATGCAGAGCAGCGCGTGGCGCCCCCAG 2090

Qy        397 ArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeuPheGluHisPheLeu 416
Db       2091 CGCGGCCTCCCCGAGCGCTCCGTCTCCATCCTCCGGGCTTGGCTCTTCGAACACTTCCTC 2150

Qy        417 HisProTyrProLysAspSerAspLysIleMetLeuAlaLysGlnThrGlyLeuThrArg 436
Db       2151 CACCCATATCCCAAGGATTCGGACAAGATCATGCTCGCCAAGCAAACCGGCCTCACCAGG 2210

Qy        437 SerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLysProMetValGlu 456
Db       2211 AGCCAGGTTTCGAACTGGTTCATCAATGCCAGGGTGAGGCTGTGGAAGCCGATGGTGGAG 2270

Qy        457 GluMetTyrLeuGluGluValLysAsnGlnGluGlnAsnSerThrAsnThrSerGlyAsp 476
Db       2271 GAGATGTACCTGGAGGAGACCAAGGACCAGGACGGCGGCGGCGGCGCCGGCGCCGGCGAC 2330

Qy        477 AsnLysAsnLys----------------------------------------------- 480
```

```
                  :::|||
Db       2331 GAGGGGAGCAAGCCCGGTGGGAGCAAGGGCGGCGGCGCCGGCGTCAATGGCGGTGTGGTT 2390

Qy        481 ------------GluThrAsnIleSerAlaProAsnGluGluLysHisProIleIleThr 496
                      :::   :::      |||   :::   |||           :::
Db       2391 GACAGTGCCGCGAAGATGGACAGCAAGGCGGCGCACATGGAGAGCGGCGGCGGCGTACAT 2450

Qy        497 SerSerLeuLeuGln-AspGlyIleThrThrThrGlnAlaGluIleSerThrSerThrIl 516
               ||||||||||:::       :::||||||   :::         ||||||:::|||
Db       2451 CCGTCGCTGCTCGAGCTCGCCGGTGACCACCAGGCGCACGGCGGGGTTCTACGACGACGAC 2510

Qy        516 eSerThrSerProThrAlaGlyAlaSerLeuHisHisAlaHisAsnPheSerPheLeu-- 535
              :::             |||   |||:::|||                 :::
Db       2511 GACGAGGATGGCGGC-GCCGCCGCCGCGCTGCAGCAGAAGCTGAAGAAGGCGAGGACGGA 2569

Qy        536 ------------GlySerPheAsnMetAspAsnThrThrThrThrValAspHisIleGl 551
                         :::|||::::::   :::      |||        |||
Db       2570 GGAGCAGCAGCAGGCGGCGTTCCACGTGTCCGACGTGGCCACGCTGCACGCGCATGCCGC 2629

Qy        551 uAsnAsnAlaLysLysGlnArgAsnAsp-----------------MetHisLysPheSe 565
               |||      :::   ::::::                  :::   ||||||
Db       2630 GGCGGCGGCGGCGGCGAGGCACGACGAGGTGAGCCACCGGGAGCTCCTCATGAAGTTCAT 2689

Qy        565 rProSerSerIleLeuSerSerValAspMetGluAlaLysAlaArgGluSerSerAsnLy 585
               |||        :::             |||  ||||||:::     :::::
Db       2690 GGAGAGCGGC---AGCGCCGGCGCCGGCGCCGGCGCCGCCGCGAGGGACCACCACCACGA 2746

Qy        585 sGlyPheThrAsnProLeuMetAlaAlaTyrAlaMetGlyAspPheGlyArgPheAspPr 605
               :              :::   :::|||         :::|||:::|||
Db       2747 ACACCATGGCGGCGTCGGCTACTCGCTGTTCGCGCCGGCGCCGTACGGGCAGTTC----- 2801

Qy        605 oHisAspGlnGlnMetThrAlaAsnPheHis-----------GlyAsnAsnGlyValSe 621
               :::|||                |||           |||      ||||||||
Db       2802 -GCCACGGAGCAGTTCGCGTTCGCCGGCCACGGCGGCGGTGGCGGCGGCGGCGGCGTGTC 2860

Qy        621 rLeuThrLeuGlyLeuPro---ProSerGluAsnLeuAlaMetProValSerGlnGlnAs 640
              ||||||||||||||||||      :::|||      |||                ::
Db       2861 GCTCACGCTCGGCCTCCCGCACGGCGCCGAGCAGACGGCG-----------------TC 2902

Qy        640 nTyrLeuSerAsnAspLeuGlySerArgSerGluMetGlySerHisTyrAsnArgMetGl 660
               ::::|||          :::   |||:::      |||              ||
Db       2903 GTTCCTCATGACCAGC---AGCAACGGCAGCGACGGCGCCGGCCACGTCGCCGGCGGCGG 2959

Qy        660 yTyrGluAsnIleAspPheGlnSerGlyAsnLysArgPheProThrGlnLeuLeuProAs 680
               |      :::::::::   ||||||      |||   |||      ||||||::   ||
Db       2960 CGGCTACGACATGAACATGCAGAGC---ACAAAGTCCTTCGCGGCTCAGCTCATGAGAGA 3016

Qy        680 pPheVal 682
               |||||||
Db       3017 CTTCGTG 3023

RESULT 12
US-10-424-599-63874
; Sequence 63874, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 63874
;   LENGTH: 1892
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_28692C.1
US-10-424-599-63874
```

```
Alignment Scores:
Pred. No.:              4.89e-93      Length:          1892
Score:                  1037.00       Matches:         254
Percent Similarity:     62.88%        Conservative:    56
Best Local Similarity:  51.52%        Mismatches:      108
Query Match:            29.10%        Indels:          76
DB:                     18            Gaps:            18

US-10-624-201A-2 (1-688) x US-10-424-599-63874 (1-1892)

Qy        253 AsnSerSerGlyGlyGlyGluSerSerSerArgGlnLys---AsnGluValAlaValGlu 271
              :::|||   |||||||||||:::::::::              :::      ||||||
Db         25 TCGAGCTGTGGAGGAGGGGAGAATAATAATAATAATGATGGAGGCAAACAAGGAGTTGAA 84

Qy        272 LeuThrThrAlaGlnArgGlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeu 291
              |||:::|||||||||||||||||||||||||||||:::||||||:::   ||||||
Db         85 CTCAGCACAGCACAGAGACAAGAGCTTCAGATGAAGAAGTCAAAACTTGTGACCATGCTA 144

Qy        292 GluGluValGluGlnArgTyrArgGlnTyrHisHisGlnMetGlnIleIleValLeuSer 311
              :::|||||||||||||||||||||||||||||||||||||:::::::::    |||
Db        145 GATGAGGTGGAGCAAAGGTACCGACAGTATCACCACCAAATGCAAGTTGTGATAACATCA 204

Qy        312 PheGluGlnValAlaGlyIleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAla 331
              |||||||||   ||||||:::|||:::|||||||||||||||   |||||||||
Db        205 TTTGAGCAAGCAGCGGGTGTTGGAGCTGCAAAGTCTTACACAGCCCTTGCGTTAAAGACA 264

Qy        332 IleSerLysGlnPheArgCysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSer 351
              |||||||||||||||||||||||||||||||||||:::   |||:::|||   ||||||
Db        265 ATCTCAAAGCAATTTCGGTGTCTCAAAGATGCAATCTCTTCACAAATAAAGACCACGAGC 324

Qy        352 LysSerLeuGlyGluGluGluGlyLeuGlyGlyLysIleGluGlySerArgLeuLysPhe 371
              |||:::|||||||||:::::   ||||||   |||:::|||||||||||||||:::::
Db        325 AAAACCTTAGGTGAAGATGATTGCTTGGGAGTTAAGGTAGAAGGTTCGAGGCTTAGGTAT 384

Qy        372 ValAspHisHisLeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGlnProAsn 391
              ||||||||||   ||||||||||||||||||||||||:::||||||
Db        385 GTTGATCATCAACTGAGACAACAACGTGCATTACAACAGCTTGGAATGCGGACGCGTGGG 444

Qy        392 AlaTrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeu 411
                     |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        445 CGGACGCGTGGGCAAAGAGGCTTGCCTGAACGTGCTGTTTCTGTTCTTCGAGCTTGGCTT 504

Qy        412 PheGluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGln 431
              |||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||
Db        505 TTTGAACATTTCTTGCACCCTTATCCTAAGGACTCCGATAAGGTTATGCTTGCAAAACAA 564

Qy        432 ThrGlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrp 451
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        565 ACTGGACTTACTCGGAGCCAGGTGTCTAACTGGTTTATAAATGCCCGAGTTCGGCTATGG 624

Qy        452 LysProMetValGluGluMetTyrLeuGluGluValLysAsnGlnGluGlnAsnSerThr 471
              |||||||||||||||||||||||||||||||||||||||||||   :::      ||||||:::
Db        625 AAGCCAATGGTAGAAGAAATGTACTTGGAAGAGGTAAAACAAGAACCAAACAATAGCTCT 684

Qy        472 ------AsnThrSer-GlyAspAsnLysAsnLysGluThrAsnIleSerAlaProAsnGl 489
                    ||||||:::    |||   ||||||:::      :::
Db        685 CAGGATAACACCACAAAAAGATCAAAAGAATCAAGCAAGGAGTTATGGT---------- 733

Qy        489 uGluLysHisProIleIleThrSerSerLeuLeuGln---AspGlyIleThrThrThr-- 507
               :::|||      ::::::      :::|||:::|||   |||   ||||||:::|||
Db        734 -CAGAAGCTAATGCTACTGCACAGGAATCTGGTGCAATGAGACTTGATCACATCAACATT 792

Qy        508 ----------GlnAlaGluIleSerThrSerThrIleSerThrSerPro---ThrAlaGl 523
                         :::   :::   ||||||:::|||   :::|||   |||
Db        793 CTCCAATCAAAAGCAGAAAGCTTCAACAACAACAACAACAACCAAACCACTTCCCCAACA 852

Qy        523 yAlaSerLeuHisHisAlaHisAsnPheSerPheLeuGlySerPheAsnMetAspAsnTh 543
               ||||||   ::::::||||||                     |||::::::   :::
Db        853 GAGATCTCTAAT-AATTCCCATAATTCACTCCAATCAGCTGGTTTCCATCTTGCTGACAT 911

Qy        543 rThrThrThrValAspHisIleGluAsnAsnAlaLysLysGlnArgAsnAspMetHisLy 563
                  :::            |||   :::   ::::::   |||
```

```
Db        912 GCAAAGCCCAAACAAGCCAATTAGAAGCACTTCTGAAATGCAGAAC-------------- 957

Qy        563 sPheSerProSerSerIleLeuSerSerValAspMetGluAlaLysAlaArg-------- 580

Db        958 ----TCTCCAGGTAGCATTCTT---TCAGTGGACATGGAAATGAAGCCTCGTCATCATGG 1010

Qy        581 -GluSerSerAsn-----------------LysGlyPheThrAsnProLeuMet----- 592

Db       1011 AGAGACAAATAATAATACTATTATTACTAGAGAAGGCAACAACAACACAACAAAGTTTGG 1070

Qy        593 -----------------------------------AlaAlaTyrAlaMetGlyAspPh 600

Db       1071 CATTGAAAGCCATGGTGGTGGTGGTGGTGGTTTTGGAGCATTTCCCAATATGGAAGACAT 1130

Qy        600 eGlyArgPheAspProHis---------------AspGlnGlnMetThrAla---AsnPh 614

Db       1131 TGGAAGGTTCCATCATCATCATCATCATGTCACTGATCAGCAATCACTGGCTCCTAGGTT 1190

Qy        614 eHisGlyAsnAsnGlyValSerLeuThrLeuGlyLeuProProSer---GluAsn----- 631

Db       1191 TCATGGA---AATGGTGTTTCTCTCACTCTTGGACTTCCACACAGCACTGAGAATAATAA 1247

Qy        632 ----------LeuAlaMetProValSerGlnGlnAsnTyrLeuSerAsnAsp------Le 646

Db       1248 TAACCTTTCCCTCTCAGGAACTACTCATCAACATGGATTCCTCTCACAGAACATGCACTT 1307

Qy        646 uGlySerArgSerGluMetGlySerHisTyrAsnArgMet------------------- 659

Db       1308 GGGAATGAGATCAACAACCAATAATGCGAATAATGAGTTTTGTGGTGCCATCAACACTAC 1367

Qy        660 ----------------------------GlyTyrGluAsnIleAspPheGlnSerGlyAs 670

Db       1368 TCCACCTTCTAATTCTCACTCAGGCACCAGTTATGAGAGCATTGACATTATTCAAAACAG 1427

Qy        670 nLysArgPheProThrGlnLeuLeuProAspPheVal 682

Db       1428 AAAGAGATTTGCTGCGCAGTTGTTGCGGGATTTCGTG 1464

RESULT 13
US-10-425-114-15195
; Sequence 15195, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 15195
;   LENGTH: 2313
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: LIB3060-003-B11_FLI
US-10-425-114-15195

Alignment Scores:
Pred. No.:              4.2e-89         Length:         2313
Score:                  999.00          Matches:        246
Percent Similarity:     56.21%          Conservative:   80
Best Local Similarity:  42.41%          Mismatches:     155
Query Match:            28.04%          Indels:         99
DB:                     18              Gaps:           18

US-10-624-201A-2 (1-688) x US-10-425-114-15195 (1-2313)

Qy          2 TyrTyrGlnGlyThrSerAspAsnThrAsnIleGlnAlaAspHisGlnGlnArgHisAsn 21

Db        675 TACTACCACGGTGGCGCC---GGCACGGACATCCAGGCC---------------------- 710

Qy         22 HisGlyAsnSerAsnAsnAsnAsnIleGlnThrLeuTyrLeuMetAsnProAsnAsnTyr 41

Db        711 -----------AGCACCGACGGCCTGCAGACGCTGTACCTCATGAACCCG---AGCTAC 755

Qy         42 MetGlnGlyTyrThrThr---------SerAspThrGlnGlnGlnGlnGlnLeuLeuPhe 58

Db        756 GCCGGCGGCTACGCTGACGACGGCGGCGCGTCCACGCCGGGGGCGGCCAACATGATGCTC 815

Qy         59 LeuAsnSerSerProAlaAla---------SerAsnAlaLeuCysHisAlaAsnIleGln 75

Db        816 CTCAACTCGGCCGTGACCACCATGACGCCGGCGTCCTTCGCGCACCACCACCAGCAGCAG 875

Qy         76 HisAlaProLeuGlnGlnGlnHisPheValGlyValProLeuProAlaValSerLeuHis 95

Db        876 TCGCCGCCGGCCGCGCAGCAGCACTTCGTCGGCATCCCGCTCCAGGTC------------ 923

Qy         96 AspGlnIleAsnHisHisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnSerGln 115

Db        923 ------------------------------------------------------------ 923

Qy        116 GlnValIleValProSerSerThrGlyValSerAlaThrSerCysGlyGlyIleThrThr 135

Db        924 CAGGCGCCGCCGCCGTCGGGCTACAACCTGTGGACCCCGGCCATGGCTGGCGGCGCCGCG 983

Qy        136 AspLeuAlaSerGlnLeuAlaPheGlnArgProIleProThrProGlnHisArgGlnGln 155

Db        984 GACATGTCGCCGCAGGCGCAGGCGCAGACTCCC----------------------GGC 1019

Qy        156 GlnGlnGlnGlnGlyGlyLeuSer------LeuSerLeuSerProGlnLeuGlnGlnGln 173

Db       1020 GCCGCGGGTGCTGCTGGGGTCAGCGCCGTGCTCAGCCTGTCGTCCCGCGAGGCGCCGCCG 1079

Qy        174 IleSerPheAsnAsnAsnIleSerSerSerSerProArgThrAsnAsnValThrIleArg 193

Db       1080 GTCACC------------GTGGCCGCGGTGGCCGCCTGCACCGACGAGGGCAAGTACCTG 1127

Qy        194 GlyThrLeuAspGlySerSerSerAsnMetValLeuGlySerLysTyrLeuLysAlaAla 213

Db       1128 GGCGTGTCGGCCACGCCGCAGGGGCAGATGGTCATGAGCTCCAAGTACCTCAAGGCCGCG 1187

Qy        214 GlnGluLeuLeuAspGluValValAsnIleValGlyLysSerIleLysGlyAspAspGln 233

Db       1188 CAGGAGCTGCTCGACGAGGTGGTCAGC---GTCAGCAAGGGCGTC------GAGGACGCT 1238

Qy        234 LysLysAspAsnSerMetAsnLysGluSerMetProLeuAla------SerAspValAsn 251

Db       1239 AAGACGGCCGCTGCCGCCGCGACCAAGAGCCTGGCCGCGGTGAAGAAGAAGGAGGACTCG 1298

Qy        252 ThrAsnSerSerGlyGlyGlyGlu---------SerSerSerArgGlnLysAsnGluVal 268

Db       1299 GAGGGCGCGTCAGGTGGCGGCACCGACGACGGTGCCGGCGCAAAGAGCGGCGGTGGTGCC 1358

Qy        269 AlaValGluLeuThrThrAlaGlnArgGlnGluLeuGlnMetLysLysAlaLysLeuLeu 288

Db       1359 GCGCCAGAGATGTCCACGGCGGAGCGGCAGGAGATGCAGATGAAGAAGAGCAAGCTTCTC 1418

Qy        289 AlaMetLeuGluGluValGluGlnArgTyrArgGlnTyrHisHisGlnMetGlnIleIle 308

Db       1419 AACATGCTTGACGAGGTGGAGCAGCGGTACCGGCAGTACCACGGGCAGATGCAGGCGGTG 1478

Qy        309 ValLeuSerPheGluGlnValAlaGlyIleGlySerAlaLysSerTyrThrGlnLeuAla 328

Db       1479 TCGTCGTCGTTCGAGGCGGCGGCGGGGGCCGGGTCGGCGCGGACGTACACGGCGCTGGCG 1538

Qy        329 LeuHisAlaIleSerLysGlnPheArgCysLeuLysAspAlaIleAlaGluGlnValLys 348

Db       1539 CTGCGCACCATCTCGCGGCAGTTCCGGTGTCTGCGGGACGCGATCGCGTCGCAGGTGCGC 1598

Qy        349 AlaThrSerLysSerLeuGlyGluGluGluGly---------LeuGlyGlyLysIleGlu 365
```

```
Db       1599 GCGGCGAGCCGGGCGCTGGGCGAGGACGCGGACGCCGCCGTCGCCGGGGGCCGCACCGTG 1658

Qy        366 GlySerArgLeuLysPheValAspHisHisLeuArgGlnGlnArgAlaLeuGlnGlnIle 385

Db       1659 GGGTCCCGCCTCCGCTACATCGACCACCAGCTCCGGCAGCAGCGCGCGCTGCAGCAGCTC 1718

Qy        386 GlyMetMetGlnProAsnAlaTrpArgProGlnArgGlyLeuProGluArgAlaValSer 405

Db       1719 GGCATGATGCAGGGCGGCGCCTGGCGGCCCCAGCGCGGCCTCCCCGAGCGCTCCGTCTCC 1778

Qy        406 ValLeuArgAlaTrpLeuPheGluHisPheLeuHisProTyrProLysAspSerAspLys 425

Db       1779 ATCCTCCGCGCCTGGCTCTTCGAGCACTTCCTGCACCCATACCCCAAGGATTCGGACAAG 1838

Qy        426 IleMetLeuAlaLysGlnThrGlyLeuThrArgSerGlnValSerAsnTrpPheIleAsn 445

Db       1839 ATCATGCTCGCCAAGCAAACCGGGCTCACCAGGAGTCAGGTGTCCAATTGGTTCATCAAT 1898

Qy        446 AlaArgValArgLeuTrpLysProMetValGluGluMetTyrLeuGluGluValLysAsn 465

Db       1899 GCGAGGGTGCGGCTGTGGAAGCCTATGGTGGAGGAGATGTACCTGGAGGAGACAAAGGAC 1958

Qy        466 GlnGluGln----------AsnSerThrAsnThrSerGlyAspAsnLysAsnLysGlu 481

Db       1959 CAGGACGCTGGAGGCGGCGGCAACGACGAGGGCAAGTCCGGTGGCGGCGGCAGCAAGAGC 2018

Qy        482 ThrAsnIleSerAlaProAsnGluGluLysHisProIleIleThrSerSerLeuLeuGln 501

Db       2019 AGCGACACCGTCGACGGCGTCATGCCGAGGGCGGACGTGATGTCCAAGTCCGCGGTGCGC 2078

Qy        502 AspGly----------IleThrThr-ThrGlnAlaGluIle----- 511

Db       2079 GTGGGAGGCGGCGGGGCGGGGGCGGAGAGCGCGTCCACCAACAAGGGCATCCATGGCTCC 2138

Qy        512 ----------SerThrSerThrIleSerThrSerProThrAlaGlyAlaSerLeuHi 527

Db       2139 TCCCTGCTCGAGCTCGGCGGCGGCGATCACCAGCAGTCCCACGCGGGGTTCTACGACG-- 2196

Qy        527 sHisAlaHisAsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrThr 546

Db       2197 ----------ACGATGAAGACGACGGCGACGACG 2220

RESULT 14
US-10-425-115-97705
; Sequence 97705, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(53222)B
;  CURRENT APPLICATION NUMBER: US/10/425,115
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 97705
;   LENGTH: 2970
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_20618C.1
US-10-425-115-97705

Alignment Scores:
Pred. No.:              6.11e-89      Length:          2970
Score:                  999.00        Matches:         246
Percent Similarity:     56.21%        Conservative:    80
Best Local Similarity:  42.41%        Mismatches:      155
Query Match:            28.04%        Indels:          99
DB:                     20            Gaps:            18

US-10-624-201A-2 (1-688) x US-10-425-115-97705 (1-2970)

Qy          2 TyrTyrGlnGlyThrSerAspAsnThrAsnIleGlnAlaAspHisGlnGlnArgHisAsn 21

Db       1034 TACTACCACGGTGGCGCC---GGCACGGACATCCAGGCC---------- 1069

Qy         22 HisGlyAsnSerAsnAsnAsnAsnIleGlnThrLeuTyrLeuMetAsnProAsnAsnTyr 41

Db       1070 -----------AGCACCGACGGCCTGCAGACGCTGTACCTCATGAACCCG---AGCTAC 1114

Qy         42 MetGlnGlyTyrThrThr---------SerAspThrGlnGlnGlnGlnGlnLeuLeuPhe 58

Db       1115 GCCGGCGGCTACGCTGACGACGGCGGCGCGTCCACGCCGGGGGCGGCCAACATGATGCTC 1174

Qy         59 LeuAsnSerSerProAlaAla---------SerAsnAlaLeuCysHisAlaAsnIleGln 75

Db       1175 CTCAACTCGGCCGTGACCACCATGACGCCGGCGTCCTTCGCGCACCACCACCAGCAGCAG 1234

Qy         76 HisAlaProLeuGlnGlnGlnHisPheValGlyValProLeuProAlaValSerLeuHis 95

Db       1235 TCGCCGCCGGCCGCGCAGCAGCACTTCGTCGGCATCCCGCTCCAGGTC---------- 1282

Qy         96 AspGlnIleAsnHisHisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnSerGln 115

Db       1282 ---------- 1282

Qy        116 GlnValIleValProSerSerThrGlyValSerAlaThrSerCysGlyGlyIleThrThr 135

Db       1283 CAGGCGCCGCCGCCGTCGGGCTACAACCTGTGGACCCCGGCCATGGCTGGCGGCGCCGCG 1342

Qy        136 AspLeuAlaSerGlnLeuAlaPheGlnArgProIleProThrProGlnHisArgGlnGln 155

Db       1343 GACATGTCGCCGCAGGCGCAGGCGCAGACTCCC----------GGC 1378

Qy        156 GlnGlnGlnGlnGlyGlyLeuSer------LeuSerLeuSerProGlnLeuGlnGlnGln 173

Db       1379 GCCGCGGGTGCTGCTGGGGTCAGCGCCGTGCTCAGCCTGTCGTCCCGCGAGGCGCCGCCG 1438

Qy        174 IleSerPheAsnAsnAsnIleSerSerSerSerProArgThrAsnAsnValThrIleArg 193

Db       1439 GTCACC----------GTGGCCGCGGTGGCCGCCTGCACCGACGAGGGCAAGTACCTG 1486

Qy        194 GlyThrLeuAspGlySerSerSerAsnMetValLeuGlySerLysTyrLeuLysAlaAla 213

Db       1487 GGCGTGTCGGCCACGCCGCAGGGGCAGATGGTCATGAGCTCCAAGTACCTCAAGGCCGCG 1546

Qy        214 GlnGluLeuLeuAspGluValValAsnIleValGlyLysSerIleLysGlyAspAspGln 233

Db       1547 CAGGAGCTGCTCGACGAGGTGGTCAGC---GTCAGCAAGGGCGTC------GAGGACGCT 1597

Qy        234 LysLysAspAsnSerMetAsnLysGluSerMetProLeuAla------SerAspValAsn 251

Db       1598 AAGACGGCCGCTGCCGCCGCGACCAAGAGCCTGGCCGCGGTGAAGAAGAAGGAGGACTCG 1657

Qy        252 ThrAsnSerSerGlyGlyGlyGlu---------SerSerSerArgGlnLysAsnGluVal 268

Db       1658 GAGGGCGCGTCAGGTGGCGGCACCGACGACGGTGCCGGCGCAAAGAGCGGCGGTGGTGCC 1717

Qy        269 AlaValGluLeuThrThrAlaGlnArgGlnGluLeuGlnMetLysLysAlaLysLeuLeu 288

Db       1718 GCGCCAGAGATGTCCACGGCGGAGCGGCAGGAGATGCAGATGAAGAAGAGCAAGCTTCTC 1777

Qy        289 AlaMetLeuGluGluValGluGlnArgTyrArgGlnTyrHisHisGlnMetGlnIleIle 308

Db       1778 AACATGCTTGACGAGGTGGAGCAGCGGTACCGGCAGTACCACGGGCAGATGCAGGCGGTG 1837

Qy        309 ValLeuSerPheGluGlnValAlaGlyIleGlySerAlaLysSerTyrThrGlnLeuAla 328

Db       1838 TCGTCGTCGTTCGAGGCGGCGGCGGGGGCCGGGTCGGCGCGGACGTACACGGCGCTGGCG 1897

Qy        329 LeuHisAlaIleSerLysGlnPheArgCysLeuLysAspAlaIleAlaGluGlnValLys 348
```

```
Db       1898 CTGCGCACCATCTCGCGGCAGTTCCGGTGTCTGCGGGACGCGATCGCGTCGCAGGTGCGC 1957

Qy        349 AlaThrSerLysSerLeuGlyGluGluGluGly---------LeuGlyGlyLysIleGlu 365

Db       1958 GCGGCGAGCCGGGCGCTGGGCGAGGACGCGGACGCCGCCGTCGCCGGGGGCCGCACCGTG 2017

Qy        366 GlySerArgLeuLysPheValAspHisHisLeuArgGlnGlnArgAlaLeuGlnGlnIle 385

Db       2018 GGGTCCCGCCTCCGCTACATCGACCACCAGCTCCGGCAGCAGCGCGCGCTGCAGCAGCTC 2077

Qy        386 GlyMetMetGlnProAsnAlaTrpArgProGlnArgGlyLeuProGluArgAlaValSer 405

Db       2078 GGCATGATGCAGGGCGGCGCCTGGCGGCCCCAGCGCGGCCTCCCCGAGCGCTCCGTCTCC 2137

Qy        406 ValLeuArgAlaTrpLeuPheGluHisPheLeuHisProTyrProLysAspSerAspLys 425

Db       2138 ATCCTCCGCGCCTGGCTCTTCGAGCACTTCCTGCACCCATACCCCAAGGATTCGGACAAG 2197

Qy        426 IleMetLeuAlaLysGlnThrGlyLeuThrArgSerGlnValSerAsnTrpPheIleAsn 445

Db       2198 ATCATGCTCGCCAAGCAAACCGGGCTCACCAGGAGTCAGGTGTCCAATTGGTTCATCAAT 2257

Qy        446 AlaArgValArgLeuTrpLysProMetValGluGluMetTyrLeuGluGluValLysAsn 465

Db       2258 GCGAGGGTGCGGCTGTGGAAGCCTATGGTGGAGGAGATGTACCTGGAGGAGACAAAGGAC 2317

Qy        466 GlnGluGln------------AsnSerThrAsnThrSerGlyAspAsnLysAsnLysGlu 481

Db       2318 CAGGACGCTGGAGGCGGCGGCAACGACGAGGGCAAGTCCGGTGGCGGCGGCAGCAAGAGC 2377

Qy        482 ThrAsnIleSerAlaProAsnGluGluLysHisProIleIleThrSerSerLeuLeuGln 501

Db       2378 AGCGACACCGTCGACGGCGTCATGCCGAGGGCGGACGTGATGTCCAAGTCCGCGGTGCGC 2437

Qy        502 AspGly-----------------------IleThrThr-ThrGlnAlaGluIle----- 511

Db       2438 GTGGGAGGCGGCGGGGCGGGGGCGGAGAGCGCGTCCACCAACAAGGGCATCCATGGCTCC 2497

Qy        512 -------------SerThrSerThrIleSerThrSerProThrAlaGlyAlaSerLeuHi 527

Db       2498 TCCCTGCTCGAGCTCGGCGGCGGCGATCACCAGCAGTCCCACGCGGGGTTCTACGACG-- 2555

Qy        527 sHisAlaHisAsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrThr 546

Db       2556 ---------------------------------ACGATGAAGACGACGGCGACGACG 2579

RESULT 15
US-10-425-114-9882
; Sequence 9882, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Jingdong
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Screen, Steven E
;  APPLICANT:  Tabaska, Jack E
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53313)B
;  CURRENT APPLICATION NUMBER: US/10/425,114
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 9882
;   LENGTH: 1716
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: 700875318_FLI
US-10-425-114-9882

Alignment Scores:
Pred. No.:              8.51e-88       Length:        1716
```

```
Score:                  984.00         Matches:       230
Percent Similarity:     63.72%         Conservative:  44
Best Local Similarity:  53.49%         Mismatches:    90
Query Match:            27.62%         Indels:        66
DB:                     18             Gaps:          14

US-10-624-201A-2 (1-688) x US-10-425-114-9882 (1-1716)

Qy        306 GlnIleIleValLeuSerPheGluGlnValAlaGlyIleGlySerAlaLysSerTyrThr 325

Db          2 CAAATTGTGGTATCATCATTTGAGCAAGCAGCGGGTTATGGGGCGGCAAAATCTTACACT 61

Qy        326 GlnLeuAlaLeuHisAlaIleSerLysGlnPheArgCysLeuLysAspAlaIleAlaGlu 345

Db         62 GCCCTTGCCCTAAAAACTATCTCAAAGCAATTCAGGTGCCTGAAAGATGCAATCTCTGCA 121

Qy        346 GlnValLysAlaThrSerLysSerLeuGlyGluGluGluGlyLeuGlyGlyLysIleGlu 365

Db        122 CAAATCAAAGCCACGAGCAAGACATTGGGTGAAGATGATTGCTTAGGAGTTAAAGTAGAA 181

Qy        366 GlySerArgLeuLysPheValAspHisHisLeuArgGlnGlnArgAlaLeuGlnGlnIle 385

Db        182 GGGTCAAGGCTTAGGTATGTTGACCACCATCTTCGGCAACAAAGGGCACTTCAGCAACTA 241

Qy        386 GlyMetMetGlnProAsnAlaTrpArgProGlnArgGlyLeuProGluArgAlaValSer 405

Db        242 GGAATGATTCAGCCCAATGCATGGAGGCCCCAGAGAGGCTTGCCTGAACGAGCTGTTTCC 301

Qy        406 ValLeuArgAlaTrpLeuPheGluHisPheLeuHisProTyrProLysAspSerAspLys 425

Db        302 ATTCTTCGGGCTTGGCTTTTTGAGCATTTCCTTCACCCATATCCAAAGGACTCCGATAAA 361

Qy        426 IleMetLeuAlaLysGlnThrGlyLeuThrArgSerGlnValSerAsnTrpPheIleAsn 445

Db        362 GTTATGCTTGCTAAACAAACTGGACTTTCTAGGAGCCAGGTGTCAAACTGGTTTATCAAT 421

Qy        446 AlaArgValArgLeuTrpLysProMetValGluGluMetTyrLeuGluGluValLysAsn 465

Db        422 GCCCGAGTTCGGCTTTGGAAGCCAATGGTTGAAGAAATGTACTTGGAAGAAATCAAGGAA 481

Qy        466 GlnGluGlnAsn------SerThrAsnThrSerGlyAspAsnLysAsnLysGluThrAsn 483

Db        482 CATGAACAAGGTAATGCCTCAGAAAACACCAAATCAAAGGAATCGAGCAAGGAGTTGGGT 541

Qy        484 IleSerAlaProAsnGluGluLysHisProIleIleThrSerSerLeuLeuGlnAsp--- 502

Db        542 TCAACAGCAAATGTTGCACCCGAATCAGGTGCCATTAAACTCGATCATCTCCAATCAAAG 601

Qy        503 --------------GlyIleThrThrThrGlnAlaGluIleSerThrSerThrIleSer 517

Db        602 CAAGAAAGCTTCAAACACCAAAACACTTCTCCAACCGAGATCTCAACAAACTCT---TCA 658

Qy        518 ThrSerProThrAlaGlyAlaSerLeuHisHisAlaHisAsnPheSerPheLeuGlySer 537

Db        659 ATGTCTCCA---ATGGGAGGGTCCCTTCAATCCCCCTCTGGTTTCCATCTTGCAGGATCA 715

Qy        538 PheAsnMetAspAsnThrThrThrThrValAspHisIleGluAsnAsnAlaLysLysGln 557

Db        716 TCTGACATGCAA-------------------------ATTAGGAGTCCAAAGAAACCA 748

Qy        558 ArgAsnAspMetHisLysPheSerProSerSerIleLeuSerSerValAspMetGlu--- 576

Db        749 AGAAGTTCGGAAATTCAAAACTCTCCAAGTAGTATACTC---TCAGTGGGCATGGAAATG 805

Qy        577 ----------------------AlaLysAlaArgGluSerSerAsnLysGlyPheThr 588

Db        806 AAGCATAATAATAATGGTGATCATGCAAATAATAGAGAAGCCAACAACACAAAGTTTGGC 865

Qy        589 Asn-------------------------ProLeuMet-------------------- 592

Db        866 AATGAAAGGCACCACCAGAAAGATGGATACCCTTTGATGACTAGCAATGCAAACCATGGT 925

Qy        593 ---------AlaAlaTyrAlaMetGlyAspPheGly---ArgPheAspProHisAspGln 608
```

```
                     |||:::   |||   |||   |||   ||||||:::         :::
Db        926 GGAGGGTTTGGAGCGTTCACCATGGAAGACATGGGAAGCAGATTTAATGTCACTACTGAG 985

Qy        609 GlnMetThrAlaAsnPheHisGlyAsnAsnGlyValSerLeuThrLeuGlyLeuProPro 628
              |||:::   :::   |||||||||   ||||||||||||||||||||||||||||||
Db        986 CAACTGGCTTCAAGGTTCCATGGA---AATGGTGTTTCTCTCACACTTGGACTTCCCCAC 1042

Qy        629 SerGluAsnLeuAlaMetProValSerGlnGlnAsn---TyrLeuSerAsnAsp------ 645
              :::|||||||||:::||||||   :::|||||||:::   :::||||||   :::
Db       1043 AATGAGAACCTTTCTATGCCAGGAACTCAACAACATGGATTTCTCTCACAGAATATACAC 1102

Qy        646 LeuGlySerArgSerGluMetGlySerHisTyrAsnArgMet---------------- 659
              ||||||   |||   |||||||||::::::   |||
Db       1103 TTGGGAAGGAGACTTGAAATGGGAACAAATGGAAATGAGTTTTGTGCTATCAACACTCCA 1162

Qy        660 --------------------GlyTyrGluAsnIleAspPheGlnSerGlyAsnLysArg 672
                                  ||||||:::||||||   |||:::      :::
Db       1163 CCTTCTTCTCACTCAGGAACCACTTATGAGAGCATTGACATTCAAAACAGAAAGAGGTTT 1222

Qy        673 PheProThrGlnLeuLeuProAspPheVal 682
              |||||||||||||||||||||
Db       1223 GTTGCTCATCAGTTATTGCCAGATTTCGTG 1252
```

Search completed: July 29, 2005, 19:19:10
Job time : 1191 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 29, 2005, 14:13:12 ; Search time 5760 Seconds
(without alignments)
4546.566 Million cell updates/sec

Title: US-10-624-201A-2
Perfect score: 3563
Sequence: 1 MYYQGTSDNTNIQADHQQRH.....GNKRFPTQLLPDFVTGNLGT 688

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10624201/runat_27072005_122958_10917/app_query.fasta_1.839
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10624201_@CGN_1_1_4352_@runat_27072005_122958_10917 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1149	32.2	685	2	AW034617	AW034617 EST278301
2	1102	30.9	756	4	BG593861	BG593861 EST492539
3	1057	29.7	964	7	CK262102	CK262102 EST708180
4	1043.5	29.3	644	2	AW930221	AW930221 EST340678
5	1003.5	28.2	630	4	BG129304	BG129304 EST474950
6	981.5	27.5	610	7	CO909070	CO909070 BJ02039F0
7	966	27.1	699	5	BQ119770	BQ119770 EST605346
8	963.5	27.0	1965	9	CL978605	CL978605 OsIFCC032
9	961	27.0	708	4	BI920940	BI920940 EST540875
10	930	26.1	623	2	AW441343	AW441343 EST310739
11	920	25.8	553	1	AI773253	AI773253 EST254353
c 12	898	25.2	827	7	CV241175	CV241175 WS02511.B
13	895	25.1	789	4	BI921008	BI921008 EST540943
14	892.5	25.0	850	6	CB292855	CB292855 UCRCS01_0
15	889.5	25.0	752	4	BM535737	BM535737 EST588759
16	884.5	24.8	727	4	BI176426	BI176426 EST517371
17	875	24.6	783	2	AW688195	AW688195 NF004E06S
18	874	24.5	669	5	BQ148285	BQ148285 NF065F11F
19	871	24.4	740	7	CO111273	CO111273 GR__Eb004
20	870	24.4	568	2	AW441354	AW441354 EST310750
21	866.5	24.3	844	7	CO101205	CO101205 GR__Eb002
22	864	24.2	667	1	AI485358	AI485358 EST243679
23	839.5	23.6	1803	9	CL961014	CL961014 OsIFCC005
24	839	23.5	554	1	AI897675	AI897675 EST267118
25	835	23.4	781	7	CO098143	CO098143 GR__Ea22F
26	834	23.4	1064	7	CV477853	CV477853 57798.1 D
27	833	23.4	495	2	BE460264	BE460264 EST411755
28	829.5	23.3	777	6	CB343619	CB343619 CA32EN000
29	829.5	23.3	1954	3	CNS0ADW7	BX813490 Arabidops
30	823	23.1	523	1	AI898911	AI898911 EST268354
31	823	23.1	530	1	AI487525	AI487525 EST245847
32	823	23.1	645	4	BG887850	BG887850 EST513701
33	816	22.9	480	5	BQ509372	BQ509372 EST616787
34	801	22.5	636	4	BG887971	BG887971 EST513822
35	799.5	22.4	889	7	CO366367	CO366367 RTK1_27_A
36	795	22.3	576	1	AI894617	AI894617 EST264060
37	793.5	22.3	636	2	AW442048	AW442048 EST311444
38	789	22.1	537	5	BM954132	BM954132 sam70e12.
39	781	21.9	572	5	BP034454	BP034454 BP034454
40	779	21.9	1278	9	CL980113	CL980113 OsIFCC034
41	778	21.8	645	5	BU043836	BU043836 PP_LEa001
42	777	21.8	555	4	BM142730	BM142730 saj54b05.
43	775.5	21.8	650	4	BG457258	BG457258 NF101F03P
44	774	21.7	616	6	CA839352	CA839352 MCT026E10
45	768.5	21.6	747	6	CB671354	CB671354 OSJNEe04O

ALIGNMENTS

RESULT 1
AW034617
LOCUS AW034617 685 bp mRNA linear EST 18-MAY-2001
DEFINITION EST278301 tomato callus, TAMU Lycopersicon esculentum cDNA clone cLEC33H16, mRNA sequence.
ACCESSION AW034617
VERSION AW034617.1 GI:5893373
KEYWORDS EST.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 685)
AUTHORS Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T., Holt,I.E., Liang,F., Upton,J., Craven,M.B., Bowman,C.L., Ahn,S., Ronning,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
TITLE Generation of ESTs from tomato callus tissue
JOURNAL Unpublished (1999)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.
FEATURES Location/Qualifiers
source 1. .685
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEC33H16"

```
                     /tissue_type="callus"
                     /dev_stage="25-40 days old"
                     /lab_host="XL1-Blue MRF'"
                     /clone_lib="tomato callus, TAMU"
                     /note="Vector: pBlueScript SK(-); Site_1: EcoR1; Site_2:
                     Xho1; supplier: Giovannoni laboratory; cLEC - Cotyledons
                     of seedlings 7-10 days post-germination were excised, cut
                     at both ends and placed on MS medium with no selection.
                     Mixed callus was harvested at 25 and 40 days and included
                     undifferentiated masses. Tomato Callus EST Library"
ORIGIN

Alignment Scores:
Pred. No.:                1.14e-107      Length:         685
Score:                    1149.00        Matches:        222
Percent Similarity:       98.70%         Conservative:   5
Best Local Similarity:    96.52%         Mismatches:     1
Query Match:              32.25%         Indels:         2
DB:                       2              Gaps:           1

US-10-624-201A-2 (1-688) x AW034617 (1-685)

Qy        321 AlaLysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArgCysLeuLys 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCAAAATCATACACTCAATTAGCTTTGCATGCAATTTCAAAGCAATTCAGATGCCTAAAG 60

Qy        341 AspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGluGluGluGlyLeu 360
              |||||||||:::|||||||||||||||||||||||||||||||||:::|||||||||
Db         61 GATGCAATTAGTGAGCAAGTGAAGGCGACGAGCAAGAGTTTAGGTGAAGATGAAGGCTTG 120

Qy        361 GlyGlyLysIleGluGlySerArgLeuLysPheValAspHisHisLeuArgGlnGlnArg 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGAGGGAAAATTGAAGGCTCAAGACTCAAATTTGTGGACCATCATCTAAGGCAACAACGC 180

Qy        381 AlaLeuGlnGlnIleGlyMetMetGlnProAsnAlaTrpArgProGlnArgGlyLeuPro 400
              ||||||||||||:::||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCGCTGCAACAGCTAGGAATGATGCAACCAAATGCTTGGAGACCCCAAAGAGGTTTACCT 240

Qy        401 GluArgAlaValSerValLeuArgAlaTrpLeuPheGluHisPheLeuHisProTyrPro 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTTTTCGAGCATTTTCTTCATCCTTACCCC 300

Qy        421 LysAspSerAspLysIleMetLeuAlaLysGlnThrGlyLeuThrArgSerGlnValSer 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAAGACTCAGACAAAATCATGCTTGCTAAGCAAACGGGGCTAACAAGGAGCCAGGTCTCT 360

Qy        441 AsnTrpPheIleAsnAlaArgValArgLeuTrpLysProMetValGluGluMetTyrLeu 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AACTGGTTCATTAATGCTCGAGTTCGATTATGGAAGCCAATGGTTGAAGAAATGTACTTG 420

Qy        461 GluGluValLysAsnGlnGluGlnAsnSerThrAsnThrSerGlyAspAsnLysAsnLys 480
              ||||||||||||||||||||||||||||||:::|||||||||||||||||||||||||||
Db        421 GAAGAAGTGAAGAATCAAGAACAAAACAGTAGTAATACTTCAGGAGATAACAAAAACAAA 480

Qy        481 GluThrAsnIleSerAlaProAsnGluGluLysHisProIleIleThrSerSerLeuLeu 500
              ||||||||||||||||||||||||||||||||   ||||||||||||||||||||||||||
Db        481 GAGACGAATATAAGTGCTCCAAATGAAGAGAAACAACCAATTATTACTAGCAGCTTATTA 540

Qy        501 GlnAspGlyIleThrThrThrGlnAlaGluIleSerThrSerThrIleSerThrSerPro 520
              |||||||||      ||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CAAGATGGT------ACTACTCAAGCAGAAATTTCTACCTCAACTATTTCAACTTCCCCT 594

Qy        521 ThrAlaGlyAlaSerLeuHisHisAlaHisAsnPheSerPheLeuGlySerPheAsnMet 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 ACTGCAGGTGCTTCACTTCATCATGCTCACAACTTCTCCTTCCTTGGTTCATTCAACATG 654

Qy        541 AspAsnThrThrThrThrValAspHisIle 550
              :::|||||||||||||||||||||||||||
Db        655 GAGAATACTACTACTACTGTTGATCATATT 684

RESULT 2
BG593861
LOCUS       BG593861                 756 bp    mRNA    linear   EST 07-MAR-2003
DEFINITION  EST492539 cSTS Solanum tuberosum cDNA clone cSTS5P9 5' sequence,
            mRNA sequence.
ACCESSION   BG593861
VERSION     BG593861.1  GI:13612001
KEYWORDS    EST.
SOURCE      Solanum tuberosum (potato)
  ORGANISM  Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum.
REFERENCE   1  (bases 1 to 756)
  AUTHORS   van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Chiemingo,A.,
            Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.
  TITLE     Generations of ESTs from sprouting potato eyes
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robin Buell
            The Institute for Genomic Research
            9712 Medical Center Dr, Rockville, MD 20850, USA
            Email: potato-array@tigr.org
            This clone can be obtained from the University of Arizona Genomics
            Institute. Orders can be made through URL:
            http://genome.arizona.edu/orders/
            Seq primer: M13F-R.
FEATURES             Location/Qualifiers
     source          1. .756
                     /organism="Solanum tuberosum"
                     /mol_type="mRNA"
                     /cultivar="Kennebec"
                     /db_xref="taxon:4113"
                     /clone="cSTS5P9"
                     /tissue_type="sprouting eyes from tubers"
                     /dev_stage="12-14 weeks post harvest"
                     /lab_host="SOLR"
                     /clone_lib="cSTS"
                     /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
                     XhoI; Various sizes of sprouting eyes (2mm to 15mm) were
                     taken from tubers. The tubers were incubated at 26C in the
                     dark for 2-3 weeks prior to sprouting. The eyes were
                     frozen in liquid nitrogen immediately upon removal from
                     tubers."
ORIGIN

Alignment Scores:
Pred. No.:                9.41e-103      Length:         756
Score:                    1102.00        Matches:        210
Percent Similarity:       99.06%         Conservative:   0
Best Local Similarity:    99.06%         Mismatches:     2
Query Match:              30.93%         Indels:         0
DB:                       4              Gaps:           0

US-10-624-201A-2 (1-688) x BG593861 (1-756)

Qy          1 MetTyrTyrGlnGlyThrSerAspAsnThrAsnIleGlnAlaAspHisGlnGlnArgHis 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        119 ATGTACTATCAAGGAACCTCGGATAATACTAATATACAAGCTGATCATCAACAACATCAT 178

Qy         21 AsnHisGlyAsnSerAsnAsnAsnAsnIleGlnThrLeuTyrLeuMetAsnProAsnAsn 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        179 AATCATGGGAATAGTAATAATAATAATATTCAGACACTTTATTTGATGAACCCTAACAAT 238

Qy         41 TyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGlnGlnGlnLeuLeuPheLeuAsn 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 TATATGCAAGGCTACACTACTTCTGACACACAGCAGCAGCAGCAGTTACTTTTCCTGAAT 298

Qy         61 SerSerProAlaAlaSerAsnAlaLeuCysHisAlaAsnIleGlnHisAlaProLeuGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 TCTTCACCAGCAGCAAGCAACGCGCTTTGCCATGCGAATATACAACACGCGCCGCTGCAA 358

Qy         81 GlnGlnHisPheValGlyValProLeuProAlaValSerLeuHisAspGlnIleAsnHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 CAGCAGCACTTTGTCGGTGTGCCTCTTCCGGCAGTAAGTTTGCACGATCAGATCAATCAT 418

Qy        101 HisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnSerGlnGlnValIleValPro 120
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        419 CATGGACTTTTACAGCGCATGTGGAACAACCAAGATCAATCTCAGCAGGTGATAGTACCA 478

Qy        121 SerSerThrGlyValSerAlaThrSerCysGlyGlyIleThrThrAspLeuAlaSerGln 140
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        479 TCGTCGACGGGGGTTTCTGCCACGTCATGTGGCGGGATCACCACGGACTTGGCGTCTCAA 538

Qy        141 LeuAlaPheGlnArgProIleProThrProGlnHisArgGlnGlnGlnGlnGlnGlyGly 160
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        539 TTGGCGTTTCAGAGGCCGATTCCGACACCACAACACCGACAGCAGCAACAACAGCAAGGC 598

Qy        161 GlyLeuSerLeuSerLeuSerProGlnLeuGlnGlnGlnIleSerPheAsnAsnAsnIle 180
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        599 GGTCTATCTCTAAGCCTTTCTCCTCAGCTACAACAGCAAATTAGTTTCAATAACAATATT 658

Qy        181 SerSerSerSerProArgThrAsnAsnValThrIleArgGlyThrLeuAspGlySerSer 200
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        659 TCATCCTCATCACCAAGGACAAATAATGTTACTATTAGGGGAACATTAGATGGGAAGTCT 718

Qy        201 SerAsnMetValLeuGlySerLysTyrLeuLysAla 212
                 ||||||||||||||||||||||||||||||||||||
Db        719 AGCAACATGGTTTTAGGCTCTAAGTATCTGAAAGCT 754
```

RESULT 3
CK262102
LOCUS CK262102 964 bp mRNA linear EST 03-AUG-2004
DEFINITION EST708180 potato abiotic stress cDNA library Solanum tuberosum cDNA clone POABE14 5' end, mRNA sequence.
ACCESSION CK262102
VERSION CK262102.1 GI:39819080
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiids; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 964)
AUTHORS Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.
TITLE Generation of ESTs from abiotic stressed potato tissue
JOURNAL Unpublished (2003)
COMMENT Other_ESTs: EST708181
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from the University of Arizona Genomics Institute via http://genome.arizona.edu/orders/ .
Seq primer: ATT TAG GTG ACA CTA TAG.
FEATURES Location/Qualifiers
source 1. .964
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POABE14"
/tissue_type="abiotic stress treated leaf and root tissue"
/lab_host="DH10B-TonA"
/clone_lib="potato abiotic stress cDNA library"
/note="Vector: pCMVSport6.1; Site_1: EcoRI; Site_2: NotI; supplier: Solanum tuberosum var. Kennebec plants were grown from cuttings on a 16hr light/8 hr dark cycle at 25 C for 3-4 weeks. Abiotic stress conditions were applied to four separate sets of plants. Set 1 involved saturation of the soil with 150 mM NaCl and tissues were harvested at following application of the salt stress (leaves: 2hr, 6hr, 12hr, 1d, 2d, and 4d; roots:2hr, 6hr, 12hr, and 2d). Set 2 were grown under the standard conditions and then were water stressed by withdrawal of further watering applications. Drought stressed plants were harvested after cessation of watering (leaves: 3d, 5d, and 7d; roots:3d and 5d). Set 3 were grown under the standard conditions and then were cold stressed by placement at 4 C. Cold stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d. Set 4 were grown under the standard conditions and then were heat stressed by placement at 35 C. Heat stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d and 4d and heat-stressed roots were harvested at 6 hr, 12 hr, 1 d, and 4d. RNA was isolated from all tissues and equal RNA from each tissue and stress was pooled to construct the cDNA library. RNA sample."

ORIGIN

Alignment Scores:

Pred. No.:	6.08e-98	Length:	964
Score:	1057.00	Matches:	226
Percent Similarity:	77.35%	Conservative:	37
Best Local Similarity:	66.47%	Mismatches:	51
Query Match:	29.67%	Indels:	27
DB:	7	Gaps:	9

US-10-624-201A-2 (1-688) x CK262102 (1-964)

```
Qy        237 AsnSerMetAsnLysGluSerMetProLeuAlaSerAspValAsnThrAsnSerSerGly 256
                 :::   ::::::   |||   :::   ||||||||||||
Db          8 CACAAATTGGACAATGAATTAATATCTTTGGCTAGTGAT---------------------- 46

Qy        257 GlyGlyGluSerSerSerArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGln 276
                       ||||||   |||||||||   :::   ||||||||||||||||||||||
Db         47 ---GATGTTGAAAGTAGCAGCCAAAAAAAT---ATTGTTGTTGAACTTACTACAGCTCAA 100

Qy        277 ArgGlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeuGluGluValGluGln 296
                 ||||||||||||||||||||||||||||||||||:::::|||||:::|||||::|||
Db        101 AGACAAGAACTTCAAATGAAGAAAGCCAAGCTTGTTAGCATGCTTGATGAGGTGGATCAA 160

Qy        297 ArgTyrArgGlnTyrHisHisGlnMetGlnIleIleValLeuSerPheGluGlnValAla 316
                 ||||||||||||||||||||||||||||||:::|||       ||||||||||||
Db        161 AGGTATAGACAATACCATCACCAAATGCAAATGATTGCAACATCATTTGAGCAAACAACA 220

Qy        317 GlyIleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPhe 336
                 ||||||||||||:::|||||||||||||||||||||||||   ||||||||||||||
Db        221 GGAATTGGATCATCAAAATCATACACACAACTTGCTTTGCACACAATTTCGAAGCAATTT 280

Qy        337 ArgCysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGlu 356
                 ||||||||||||||||||:::   |||:::|||   |||||||||:::||||||||||
Db        281 AGATGTTTAAAAGATGCAATTTCTGGGCAAATAAAGGACACAAGTAAAACTTTAGGGGAA 340

Qy        357 GluGluGlyLeuGlyGlyLysIleGluGlySerArgLeuLysPheValAspHisHisLeu 376
                 ||||||   :::|||||||||||||||||:::||||||||||||||||||||||||||||
Db        341 GAAGAGAACATTGGAGGCAAAATTGAAGGATCAAAGTTGAAATTTGTGGATCATCATTTA 400

Qy        377 ArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGlnProAsnAlaTrpArgProGln 396
                 ||||||||||||||||||||:::||||||||||||   ||||||||||||||||||
Db        401 CGCCAACAACGTGCACTACAACAATTAGGGATGATGCAAACCAATGCATGGAGGCCACAG 460

Qy        397 ArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeuPheGluHisPheLeu 416
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 AGAGGTTTGCCCGAAAGAGCGGTTTCGGTTCTCCGCGCTTGGCTTTTCGAGCATTTTCTT 520

Qy        417 HisProTyrProLysAspSerAspLysIleMetLeuAlaLysGlnThrGlyLeuThrArg 436
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        521 CATCCGTATCCCAAAGATTCAGATAAAATCATGCTTGCTAAGCAAACAGGGCTAACAAGG 580

Qy        437 SerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLysProMetValGlu 456
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        581 AGCCAGGTATCAAATTGGTTTATAAATGCTAGAGTTAGACTATGGAAGCCAATGGTAGAA 640

Qy        457 GluMetTyrLeuGluGluValLys------AsnGlnGluGlnAsnSerThrAsnThrSer 474
                 |||||||||:::|||||||||||||      |||||||||||||||          :::
Db        641 GAAATGTACATGGAAGAAGTGAAGAAAACCAATCAAGAACAAAATATTGAGCCTAATAAC 700

Qy        475 GlyAspAsnLysAsnLysGluThrAsnIleSerAlaProAsnGluGluLysHisProIle 494
                 :::             ::::::::::::            :::||||||   |||
Db        701 AATGAAATTGTTGGTTCAAAATCAAGTGTTCCA---------CAAGAGAAATTACCA--- 748

Qy        495 IleThrSerSerLeuLeuGlnAspGlyIleThrThrThrGlnAlaGluIleSerThrSer 514
```

```
              |||:::|||:::::::::   :::          :::      :::|||||||||||||
Db        749 ATTAGTAGCAATATTATTCATAAT---------GCTTCTCCAAATGATATTTCTACTTCC 799

Qy        515 ThrIleSerThrSerProThrAlaGlyAlaSerLeuHisHisAlaHis---------Asn 531
              |||||||||||||||||||   |||              |||
Db        800 ACCATTTCAACATCTCCGACAGGCGGCGGCGGTTCGATTCCGGCTCAGACGGTTGCAGGT 859

Qy        532 PheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrValAspHisIleGlu 551
              |||||||||:::   |||   ||||||:::|||          :::|||
Db        860 TTCTCCTTCATTAGGTCATTAAACATGGAGAAC------------ATTGATGATCAAAGG 907

Qy        552 AsnAsnAlaLysLysGlnArgAsnAspMetHisLysPheSerProSerSerIleLeuSer 571
              ||||||   |||||||   ||||||:::|||   |||         |||:::||||||||||
Db        908 AACAAC---AAAAAGGCAAGAAATGAGATGCA-AAATTGTTCACTAGTACTATTCTCTCA 963

RESULT 4
AW930221
LOCUS       AW930221                644 bp    mRNA    linear   EST 18-MAY-2001
DEFINITION  EST340678 tomato fruit mature green, TAMU Lycopersicon esculentum
            cDNA clone cLEF41L1 5', mRNA sequence.
ACCESSION   AW930221
VERSION     AW930221.1  GI:8105622
KEYWORDS    EST.
SOURCE      Lycopersicon esculentum (tomato)
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1  (bases 1 to 644)
  AUTHORS   Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F.,
            Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S.,
            Ronning,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and
            Giovannoni,J.
  TITLE     Generation of ESTs from tomato fruit tissue
  JOURNAL   Unpublished (1999)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            5 prime sequence.
FEATURES             Location/Qualifiers
     source          1. .644
                     /organism="Lycopersicon esculentum"
                     /mol_type="mRNA"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cLEF41L1"
                     /tissue_type="fruit pericarp"
                     /dev_stage="mature green (3-5 days pre-ripening)"
                     /lab_host="SOLR"
                     /clone_lib="tomato fruit mature green, TAMU"
                     /note="Vector: pBlueScript SK(-); Site_1: EcoR1; Site_2:
                     Xho1; cLEF - Fruit were tagged at the lcm stage and
                     harvested 3-5 days prior to  ripening. Fruit were cut in
                     half to verify the seeds were indeed  'immature' and the
                     seeds and locules were discarded prior to freezing the
                     pericarp"
ORIGIN

Alignment Scores:
Pred. No.:              7.95e-97        Length:         644
Score:                  1043.50         Matches:        203
Percent Similarity:     96.76%          Conservative:   6
Best Local Similarity:  93.98%          Mismatches:     4
Query Match:            29.29%          Indels:         3
DB:                     2               Gaps:           2

US-10-624-201A-2 (1-688) x AW930221 (1-644)

Qy        362 GlyLysIleGluGlySerArgLeuLysPheValAspHisHisLeuArgGlnGlnArgAla 381
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 GGGAAAATTGAAGGCTCAAGACTCAAATTTGTGGACCATCATCTAAGGCAACAACGCGCG 62

Qy        382 LeuGlnGlnIleGlyMetMetGlnProAsnAlaTrpArgProGlnArgGlyLeuProGlu 401
              |||||||||:::|||||||||||||||||||||||||||||||||||||||||||||||
Db         63 CTGCAACAGCTAGGAATGATGCAACCAAATGCTTGGAGACCCCAAAGAGGTTTACCTGAA 122

Qy        402 ArgAlaValSerValLeuArgAlaTrpLeuPheGluHisPheLeuHisProTyrProLys 421
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 AGAGCTGTCTCTGTCCTTCGTGCTTGGCTTTTCGAGCATTTTCTTCATCCTTACCCCAAA 182

Qy        422 AspSerAspLysIleMetLeuAlaLysGlnThrGlyLeuThrArgSerGlnValSerAsn 441
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 GACTCAGACAAAATCATGCTTGCTAAGCAAACGGGGCTAACAAGGAGCCAGGTCTCTAAC 242

Qy        442 TrpPheIleAsnAlaArgValArgLeuTrpLysProMetValGluGluMetTyrLeuGlu 461
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 TGGTTCATTAATGCTCGAGTTCGATTATGGAAGCCAATGGTTGAAGAAATGTACTTGGAA 302

Qy        462 GluValLysAsnGlnGluGlnAsnSerThrAsnThrSerGlyAspAsnLysAsnLysGlu 481
              ||||||||||||||||||||||||||:::|||||||||||||||||||||||||||||||
Db        303 GAAGTGAAGAATCAAGAACAAAACAGTAGTAATACTTCAGGAGATAACAAAAACAAAGAG 362

Qy        482 ThrAsnIleSerAlaProAsnGluGluLysHisProIleIleThrSerSerLeuLeuGln 501
              ||||||||||||||||||||||||||||||   ||||||||||||||||||||||||||||
Db        363 ACGAATATAAGTGCTCCAAATGAAGAGAAACAACCAATTATTACTAGCAGCTTATTACAA 422

Qy        502 AspGlyIleThrThrThrGlnAlaGluIleSerThrSerThrIleSerThrSerProThr 521
              :::|||    ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        423 GAAGGT------ACTACTCAAGCAGAAATTTCTACCTCAACTATTTCAACTTCCCCTACT 476

Qy        522 AlaGlyAlaSerLeuHisHisAlaHisAsnPheSerPheLeuGlySerPheAsnMetAsp 541
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db        477 GCAGGTGCTTCACTTCATCATGCTCACAACTTCTCCTTCCTTGGTTCATTCAACATGGAG 536

Qy        542 AsnThrThrThrThrValAspHisIleGluAsnAsnAlaLysLysGlnArgAsn---Asp 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   |||   |||
Db        537 AATACTACTACTACTGTTGATCATATTGAAAACAACGCGAAAAAGCCAAGAAATCATGAC 596

Qy        561 MetHisLysPheSerProSerSerIleLeuSerSerValAspMetGlu 576
              ||||||||||||||   |||:::|||||||||||   :::||||||
Db        597 ATGCACAAGTTTTCTTCAAGTAATATTCTTTCATCTGGTGAAATGGAA 644

RESULT 5
BG129304
LOCUS       BG129304                630 bp    mRNA    linear   EST 31-JAN-2001
DEFINITION  EST474950 tomato shoot/meristem Lycopersicon esculentum cDNA clone
            cTOF23N12 5' sequence, mRNA sequence.
ACCESSION   BG129304
VERSION     BG129304.1  GI:12629492
KEYWORDS    EST.
SOURCE      Lycopersicon esculentum (tomato)
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1  (bases 1 to 630)
  AUTHORS   van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Utterback,T.,
            Hansen,C., Ronning,C. and Tanksley,S.
  TITLE     Generation of ESTs from tomato shoot/meristem tissue
  JOURNAL   Unpublished (2001)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html.
FEATURES             Location/Qualifiers
     source          1. .630
                     /organism="Lycopersicon esculentum"
                     /mol_type="mRNA"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cTOF23N12"
                     /tissue_type="shoot/meristem"
                     /dev_stage="developing shoots from 4-6wks old plants"
```

```
                     /lab_host="SOLR"
                     /clone_lib="tomato shoot/meristem"
                     /note="Vector: pBluescript SK(-); Site_1: EcoR1; Site_2:
                     Xho1; Small expanding leaves from the growing tip were
                     taken from greenhouse plants (4-6wks old TA496). Tissue
                     was immediately frozen in liquid nitrogen."
ORIGIN

Alignment Scores:
Pred. No.:                1.03e-92        Length:         630
Score:                    1003.50         Matches:        196
Percent Similarity:       95.73%          Conservative:   6
Best Local Similarity:    92.89%          Mismatches:     6
Query Match:              28.16%          Indels:         3
DB:                       4               Gaps:           2

US-10-624-201A-2 (1-688) x BG129304 (1-630)

Qy        451 TrpLysProMetValGluGluMetTyrLeuGluGluValLysAsnGlnGluGlnAsnSer 470
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 TGGAAGCCAATGGTTGAAGAAATGTACTTGGAAGAAGTGAAGAATCAAGAACAAAACAGT 62

Qy        471 ThrAsnThrSerGlyAspAsnLysAsnLysGluThrAsnIleSerAlaProAsnGluGlu 490
              :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         63 AGTAATACTTCAGGAGATAACAAAAACAAAGAGACGAATATAAGTGCTCCAAATGAAGAG 122

Qy        491 LysHisProIleIleThrSerSerLeuLeuGlnAspGlyIleThrThrThrGlnAlaGlu 510
              |||  ||||||||||||||||||||||||||||||      ||||||||||||||||
Db        123 AAACAACCAATTATTACTAGCAGCTTATTACAAGATGGT------ACTACTCAAGCAGAA 176

Qy        511 IleSerThrSerThrIleSerThrSerProThrAlaGlyAlaSerLeuHisHisAlaHis 530
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        177 ATTTCTACCTCAACTATTTCAACTTCCCCTACTGCAGGTGCTTCACTTCATCATGCTCAC 236

Qy        531 AsnPheSerPheLeuGlySerPheAsnMetAspAsnThrThrThrThrValAspHisIle 550
              ||||||||||||||||||||||||||||||:::||||||||||||||||||||||||||||
Db        237 AACTTCTCCTTCCTTGGTTCATTCAACATGGAGAATACTACTACTACTGTTGATCATATT 296

Qy        551 GluAsnAsnAlaLysLysGlnArgAsn---AspMetHisLysPheSerProSerSerIle 569
              ||||||||||||||||||    ||||||   ||||||||||||||||||||||||||||
Db        297 GAAAACAACGCGAAAAAGCCAAGAAATCATGACATGCACAAGTTTTCTCCAAGTAGTATT 356

Qy        570 LeuSerSerValAspMetGluAlaLysAlaArgGluSerSerAsnLysGlyPheThrAsn 589
              ||||||||||:::||||||||||||||||||||||:::||||||||||||||||||||||
Db        357 CTTTCATCTGTTGAAATGGAAGCCAAAGCTAGAGAATCAACAAACAAAGGGTTTACTAAT 416

Qy        590 ProLeuMetAlaAlaTyrAlaMetGlyAspPheGlyArgPheAspProHisAspGlnGln 609
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        417 CCTTTAATGGCAGCGTACGCGATGGGAGATTTTGGAAGGTTTGATCCTCATGATCAACAA 476

Qy        610 MetThrAlaAsnPheHisGlyAsnAsnGlyValSerLeuThrLeuGlyLeuProProSer 629
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        477 ATGACCGCGAATTTTCATGGAAATAATGGTGTATCTCTTACTTTAGGACTTCCTACTTCT 536

Qy        630 GluAsnLeuAlaMetProValSerGlnGlnAsnTyrLeuSerAsnAspLeuGlySerArg 649
              ||||||||||||||||||||||||||||||||||||||||||:::|||   |||
Db        537 GAAAACCTAGCCATGCCAGTGAGCCAACAAAATTATCTTTCTAATGAGTTGCGAAGTATG 596

Qy        650 SerGluMetGlySerHisTyrAsnArgMetGly 660
              |||:::|||||||||||||||||||||||||
Db        597 CCTGAAATAGGAAGTCATTACAATAGAATGGGA 629

RESULT 6
CO909070
LOCUS       CO909070                610 bp    mRNA    linear   EST 16-AUG-2004
DEFINITION  BJ02039F01 BJ02 Capsicum annuum cDNA 5', mRNA sequence.
ACCESSION   CO909070
VERSION     CO909070.1  GI:51299373
KEYWORDS    EST.
SOURCE      Capsicum annuum
  ORGANISM  Capsicum annuum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Capsicum.
REFERENCE   1  (bases 1 to 610)
  AUTHORS   Song,W.Y., Jeon,W.B., Kim,K.S., Lee,H.H., Ko,M.K., Kim,Y.S.,
            Hong,J.C. and Oh,B.J.
  TITLE     Generation of Hot Pepper (Capsicum annuum) ESTs (Express Sequence
            Tags) from Red Ripe Fruit (Song, et al.)
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Boung-Jun Oh
            Jeonnam Biotechnology Research Center
            Namyang Bld. #603, 10-4 Gwangsan-dong, Dong-gu, Gwangju 501-180,
            Korea
            Tel: 82 62 607 2422
            Fax: 82 62 607 6205
            Email: bjo@biohub.re.kr
            Plate: 039  row: F  column: 01.
FEATURES             Location/Qualifiers
     source          1. .610
                     /organism="Capsicum annuum"
                     /mol_type="mRNA"
                     /cultivar="Nokkwang"
                     /db_xref="taxon:4072"
                     /tissue_type="red pepper fruit pericarp"
                     /clone_lib="BJ02"
                     /note="Vector: pBluescript II SK(+)XR; Site_1: EcoRI;
                     Site_2: XhoI; cDNA library was generated from red ripe
                     fruit pericarp using lambda Zap II phage vector. In vivo
                     excision was done with helper phage to generate subclone
                     in pBluescript II SK(+)XR) vector."
ORIGIN

Alignment Scores:
Pred. No.:                1.82e-90        Length:         610
Score:                    981.50          Matches:        192
Percent Similarity:       95.63%          Conservative:   5
Best Local Similarity:    93.20%          Mismatches:     6
Query Match:              27.55%          Indels:         3
DB:                       7               Gaps:           1

US-10-624-201A-2 (1-688) x CO909070 (1-610)

Qy        250 ValAsnThrAsnSerSerGlyGlyGlyGluSerSerSerArgGlnLysAsnGluValAla 269
              |||||||||:::|||         |||      |||||||||  ||||||||||||||||
Db          1 GTCAACACTAGTAGT---------GGTACTGGTAGCAGCAGGCCGAAAAATGAAGTTGCT 51

Qy        270 ValGluLeuThrThrAlaGlnArgGlnGluLeuGlnMetLysLysAlaLysLeuLeuAla 289
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db         52 GCTGAGCTTACAACTGCTCAAAGACAAGAACTTCAGATGAAAAAGGCAAAGCTTGTCACC 111

Qy        290 MetLeuGluGluValGluGlnArgTyrArgGlnTyrHisHisGlnMetGlnIleIleVal 309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        112 ATGCTTGAAGAGGTGGAACAAAGGTACAGACAGTACCATCACCAAATGCAAATAATTGTG 171

Qy        310 LeuSerPheGluGlnValAlaGlyIleGlySerAlaLysSerTyrThrGlnLeuAlaLeu 329
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        172 TCATCATTTGAGCAAGTAGCAGGAATTGGATCAGCCAAATCATACACTCAACTGGCTTTG 231

Qy        330 HisAlaIleSerLysGlnPheArgCysLeuLysAspAlaIleAlaGluGlnValLysAla 349
              ||||||||||||||||||||||||||||||||||||||||||||||||:::||||||
Db        232 CATGCAATTTCGAAGCAATTCCGATGTCTAAAAGATGCAATTGCTGAACAAATTAAGGCG 291

Qy        350 ThrSerLysSerLeuGlyGluGluGluGlyLeuGlyGlyLysIleGluGlySerArgLeu 369
              ||||||||||||||||||||||||||||:::||||||||||||||||||||||||||||
Db        292 ACGAGCAAGAGTTTAGGTGAAGAGGAAGGCATAGGAGGGAAAATTGAAGGGTCGAGACTC 351

Qy        370 LysPheValAspHisHisLeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGln 389
              ||||||||||||||||||||||||||||||||||||||||||||||:::||||||||||||
Db        352 AAATTTGTAGACCATCATCTAAGGCAACAACGTGCGCTGCAACAATTAGGAATGATGCAA 411

Qy        390 ProAsnAlaTrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAla 409
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 CCAAATGCTTGGAGACCTCAAAGAGGTCTACCTGAAAGGGCTGTCTCCGTCCTTCGTGCT 471

Qy        410 TrpLeuPheGluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAla 429
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        472 TGGCTTTTCGAGCATTTTCTTCATCCTTACCCCAAGGATTCAGACAAAATCATGCTTGCT 531

Qy        430 LysGlnThrGlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArg 449
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        532 AAGCAAACCGGACTAACAAGGAGCCAGGTCTCCAACTGGTTCATAAATGCTCGAGTTCGA 591

Qy        450 LeuTrpLysProMetVal 455
                |||||||||||||||||
Db        592 CTATGGAAGCCAATGGTG 609

RESULT 7
BQ119770
LOCUS       BQ119770                699 bp    mRNA    linear   EST 07-MAR-2003
DEFINITION  EST605346 mixed potato tissues Solanum tuberosum cDNA clone STMEM54
            5' end, mRNA sequence.
ACCESSION   BQ119770
VERSION     BQ119770.2  GI:21919664
KEYWORDS    EST.
SOURCE      Solanum tuberosum (potato)
  ORGANISM  Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum.
REFERENCE   1  (bases 1 to 699)
  AUTHORS   Buell,C.R., Hart,A., Baker,B., Tanksley,S., Fry,W., Smart,C.,
            Restrepo,S., Griffiths,H., van der Hoeven,R., Tsai,J. and
            Karamycheva,S.A.
  TITLE     Generation of a set of potato cDNA clones for microarray analyses
  JOURNAL   Unpublished (2002)
COMMENT     On Apr 17, 2002 this sequence version replaced gi:20171732.
            Contact: Robin Buell
            The Institute for Genomic Research
            9712 Medical Center Dr, Rockville, MD 20850, USA
            Email: potato-array@tigr.org
            This clone can be obtained from the University of Arizona Genomics
            Institute. Orders can be made through URL:
            http://genome.arizona.edu/orders/
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1. .699
                     /organism="Solanum tuberosum"
                     /mol_type="mRNA"
                     /cultivar="Kennebec or Binjte"
                     /db_xref="taxon:4113"
                     /clone="STMEM54"
                     /tissue_type="mixed tissues"
                     /lab_host="SOLR"
                     /clone_lib="mixed potato tissues"
                     /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
                     XhoI; supplier: Combination of untreated and Phytophthora
                     infestans-treated libraries of stolons, leaves, leaflets,
                     axillary buds of stem explants, petioles, germinating
                     eyes, tubers, or roots."
ORIGIN

Alignment Scores:
Pred. No.:              8.99e-89        Length:         699
Score:                  966.00          Matches:        183
Percent Similarity:     99.46%          Conservative:   0
Best Local Similarity:  99.46%          Mismatches:     1
Query Match:            27.11%          Indels:         0
DB:                     5               Gaps:           0

US-10-624-201A-2 (1-688) x BQ119770 (1-699)

Qy          1 MetTyrTyrGlnGlyThrSerAspAsnThrAsnIleGlnAlaAspHisGlnGlnArgHis 20
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        148 ATGTACTATCAAGGAACCTCGGATAATACTAATATACAAGCTGATCATCAACAACATCAT 207

Qy         21 AsnHisGlyAsnSerAsnAsnAsnAsnIleGlnThrLeuTyrLeuMetAsnProAsnAsn 40
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 AATCATGGGAATAGTAATAATAATAATATTCAGACACTTTATTTGATGAACCCTAACAAT 267

Qy         41 TyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGlnGlnGlnLeuLeuPheLeuAsn 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 TATATGCAAGGCTACACTACTTCTGACACACAGCAGCAGCAGCAGTTACTTTTCCTGAAT 327

Qy         61 SerSerProAlaAlaSerAsnAlaLeuCysHisAlaAsnIleGlnHisAlaProLeuGln 80
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 TCTTCACCAGCAGCAAGCAACGCGCTTTGCCATGCGAATATACAACACGCGCCGCTGCAA 387

Qy         81 GlnGlnHisPheValGlyValProLeuProAlaValSerLeuHisAspGlnIleAsnHis 100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 CAGCAGCACTTTGTCGGTGTGCCTCTTCCGGCAGTAAGTTTGCACGATCAGATCAATCAT 447

Qy        101 HisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnSerGlnGlnValIleValPro 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 CATGGACTTTTACAGCGCATGTGGAACAACCAAGATCAATCTCAGCAGGTGATAGTACCA 507

Qy        121 SerSerThrGlyValSerAlaThrSerCysGlyGlyIleThrThrAspLeuAlaSerGln 140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 TCGTCGACGGGGGTTTCTGCCACGTCATGTGGCGGGATCACCACGGACTTGGCGTCTCAA 567

Qy        141 LeuAlaPheGlnArgProIleProThrProGlnHisArgGlnGlnGlnGlnGlnGlnGly 160
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 TTGGCGTTTCAGAGGCCGATTCCGACACCACAACACCGACAGCAGCAACAACAGCAAGGC 627

Qy        161 GlyLeuSerLeuSerLeuSerProGlnLeuGlnGlnGlnIleSerPheAsnAsnAsnIle 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        628 GGTCTATCTCTAAGCCTTTCTCCTCAGCTACAACAGCAAATTAGTTTCAATAACAATATT 687

Qy        181 SerSerSerSer 184
                ||||||||||||
Db        688 TCATCCTCATCA 699

RESULT 8
CL978605
LOCUS       CL978605                1965 bp    DNA     linear   GSS 21-SEP-2004
DEFINITION  OsIFCC032180 Oryza sativa Express Library Oryza sativa (indica
            cultivar-group) genomic, genomic survey sequence.
ACCESSION   CL978605
VERSION     CL978605.1  GI:52411712
KEYWORDS    GSS.
SOURCE      Oryza sativa (indica cultivar-group)
  ORGANISM  Oryza sativa (indica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 1965)
  AUTHORS   Ma,L., Wangc,J., Chen,C., Liu,X., Su,N., Li,L., Wang,X., Cao,M.,
            Jiao,Y., Sun,N., Zhang,X., Bao,J., Sun,D., Zhao,H., Yuan,L.,
            Wong,G.K.S., Deng,X.W. and Wang,J.
  TITLE     An analysis of transcriptional regulation of the rice genome and
            its comparison to Arabidopsis
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Chen Chen
            Department of Bioinformatic
            Beijing Institute of Genomics
            Chinese Academy of Sciences, Beijing 101300, China
            Tel: 86-10-80481559
            Fax: 86-10-80488676
            Email: chenchen@genomics.org.cn
            Rice genomic sequence.
            Class: exon-trapped.
FEATURES             Location/Qualifiers
     source          1. .1965
                     /organism="Oryza sativa (indica cultivar-group)"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:39946"
                     /clone_lib="Oryza sativa Express Library"
                     /note="Oryza sativa exon trapped genomic sequences "
ORIGIN

Alignment Scores:
Pred. No.:              8.35e-88        Length:         1965
```

Score:	963.50	Matches:	281
Percent Similarity:	47.62%	Conservative:	79
Best Local Similarity:	37.17%	Mismatches:	215
Query Match:	27.04%	Indels:	181
DB:	9	Gaps:	33

US-10-624-201A-2 (1-688) x CL978605 (1-1965)

```
Qy          2 TyrTyrGlnGlyThrSerAspAsnThrAsn---IleGlnAlaAspHisGlnGlnArgHis 20
              |||:::   |||              |||:::   :::||||||
Db         10 TACTTCACAGGCGGCGGCGCCGGCACCGACGTCGTCCAGGCC----------------- 51

Qy         21 AsnHisGlyAsnSerAsnAsnAsnAsnIleGlnThrLeuTyrLeuMetAsnProAsnAsn 40
                     :::      :::   :::||||||||||||||||||||||||   :::
Db         52 -----------GCCGGCACCGACGGCCTGCAGACGCTCTACCTCATGAACCCC---AGC 96

Qy         41 TyrMetGlnGlyTyrThrThrSerAspThr--------------GlnGlnGlnGlnGln 55
              |||:::   |||:::|||   :::
Db         97 TACGTC---GGCTTCACGGACGCCGCCGCGGCGCCCGGCGGCGGCGCCGCCGCCGCCAAC 153

Qy         56 LeuLeuPheLeuAsnSerSerProAlaAlaSerAsnAlaLeuCysHisAlaAsnIle--- 74
              ::::::||||||||||||         |||   :::   |||       |||:::
Db        154 ATGGTGTTCCTCAACTCG---------GCGGTGAGCACCCTCACGCCGGCGAGCTTCGGC 204

Qy         75 ---GlnHisAlaProLeu---GlnGlnGlnHisPheValGlyValProLeuProAlaVal 92
                 |||   |||              |||||||||||||:::||||||   :::
Db        205 CACCACCACCAGCCGACACCGGCGGCGCAGCACTTCGTCGGCATCCCTCTCCAGTCGGGG 264

Qy         93 SerLeuHisAspGlnIleAsnHisHisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAsp 112
                                                    :::|||
Db        265 ------------------------------------TACAACCTGTGGGGC--------- 279

Qy        113 GlnSerGlnGlnValIleValProSerSerThrGlyValSerAlaThrSerCysGlyGly 132
                                  |||   :::||||||   :::       |||
Db        280 --------------------CCTGACGCCACCGGCGGCAACGACGTGTCG--------- 309

Qy        133 IleThrThrAspLeuAlaSerGlnLeuAlaPheGlnArgProIleProThrProGlnHis 152
                                                          |||:::|||
Db        310 ------------------------------------------------CCGCCACGCCAC 321

Qy        153 ArgGlnGlnGlnGln---------GlnGlnGlyGlyLeuSerLeuSerLeuSerProGln 169
                 |||||||||              |||      :::   ::::::||||||
Db        322 GGCGCCCAGCAGCAAGCGCCAGCAGCAGCTGGCACGTCGGCGGCGGCGGTCAGCCCCGTG 381

Qy        170 LeuGlnGlnGlnIleSerPheAsnAsnAsnIleSerSerSerSerProArgThrAsnAsn 189
              |||                                  ||||||
Db        382 CTGAGC--------------------------CTGTCGTCGCGTGAGGCGGCGCCGCCG 414

Qy        190 ValThrIleArgGlyThrLeuAspGlySerSer--------------------SerAsn 202
              ||||||:::                 ::::::
Db        415 GTCACCGTGGCCGCCGCCGCCGCCGCCGCCGTCCCCGGCGGCACGGACCAGGAGAAG 474

Qy        203 MetValLeuGlySerLysTyrLeuLysAlaAlaGlnGluLeuLeuAspGluValValAsn 222
              :::|||:::   |||:::||||||||||||||||||||||||||||||||   |||:::
Db        475 GTGGTGATGCGCTCCAGGTACCTGAAGGCGGCGCAGGAGCTGCTCGACGAGGCGGTGAGC 534

Qy        223 IleValGlyLysSerIleLysGlyAspAspGlnLysLysAspAsnSerMetAsnLysGlu 242
                 |||   |||                  :::||||||::::::|||
Db        535 ---GTCAGCAAGGGCGCCGCGACGGCGGTGAAGAAGAAGGAGGACTCGGAGGGCGGCGTG 591

Qy        243 SerMetProLeuAlaSerAspValAsnThrAsnSerSerGlyGlyGlyGluSerSerSer 262
              |||                                  ||||||   |||
Db        592 TCC-------------------------GGTGGCGGCGGCGGTGCGGAGGACGGTGGC 624

Qy        263 ArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGlnArgGlnGluLeuGlnMet 282
                 |||:::      |||   |||::::::||||||:::||||||||||||||||||
Db        625 GGCAGCAAGAGCGGCGCCGCCGCCGAGATGTCGACGGCGGAGCGGCAGGAGCTGCAGATG 684

Qy        283 LysLysAlaLysLeuLeuAlaMetLeuGluGluValGluGlnArgTyrArgGlnTyrHis 302
              ||||||:::||||||||||   ||||||:::||||||||||||||||||||||||||||||
Db        685 AAGAAGAGCAAGCTACTGAACATGCTCGATGAGGTGGAGCAGCGGTACAGGCAGTACCAC 744

Qy        303 HisGlnMetGlnIleIleValLeuSerPheGluGlnValAlaGlyIleGlySerAlaLys 322
```

```
                  |||||||||  :::      :::||||||      ||||||   |||||||||
Db        745 CGGCAGATGCAGGGGGTGGCGGCGGCGTTCGAGGCGGCGGCGGGGGCCGGGTCGGCGACG 804

Qy        323 SerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPheArgCysLeuLysAspAla 342
              :::||||||   |||||||||      ||||||:::|||||||||||||||||:::||||||
Db        805 ACGTACACGTCGCTGGCGCTGCGCACCATCTCGCGGCAGTTCCGGTGCCTCCGCGACGCC 864

Qy        343 IleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGlu-----------GluGlu 358
              ||||||   ||||||:::|||   |||:::   |||||||||           :::|||
Db        865 ATCGCGGCGCAGGTCAGGGCGGCGAGCCGGGGGCTCGGGGAGGACTGCGGCGACGACGAG 924

Qy        359 GlyLeuGlyGly---------LysIleGluGlySerArgLeuLysPheValAspHisHis 375
              |||   ||||||                  |||||||||||:::|||:::||||||
Db        925 GGCGGCGGTGGCGGCGGGAGGACGACGGTGGGGTCGCGGCTGCGGTTCATCGACCACCAG 984

Qy        376 LeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGlnProAsnAla--------- 392
              |||||||||||||||||:::||||||:::||||||:::          |||
Db        985 CTCCGGCAGCAGCGCGCGATGCAGCAGCTTGGCATGGTGCACGCCGCCGCCGCCGGTGGC 1044

Qy        393 --------------TrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeu 407
                           |||||||||||||||||||||||||||||||||||||||||||||
Db       1045 GCCGCCGGCGGCGGGTGGCGGCCGCAGCGTGGCCTCCCCGAGCGCGCCGTCTCCGTCCTC 1104

Qy        408 ArgAlaTrpLeuPheGluHisPheLeuHisProTyrProLysAspSerAspLysIleMet 427
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:::|||
Db       1105 CGCGCTTGGCTCTTCGAGCATTTCCTCCACCCATACCCCAAGGATTCGGACAAGGTCATG 1164

Qy        428 LeuAlaLysGlnThrGlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArg 447
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1165 CTCGCCAAGCAAACCGGCCTCACCAGGAGCCAGGTGTCGAACTGGTTCATCAACGCGAGG 1224

Qy        448 ValArgLeuTrpLysProMetValGluGluMetTyrLeuGluGluValLysAsnGlnGlu 467
              |||||||||||||||||||||||||||||||||   ||||||   |||   :::|||
Db       1225 GTGAGGCTGTGGAAGCCAATGGTGGAGGAGATGTACGCCGAGGAGACCAAGGCCAAGGAG 1284

Qy        468 GlnAsnSerThrAsnThrSerGlyAspAsnLysAsnLysGluThrAsnIleSerAlaPro 487
              :::
Db       1285 GAG---------------------------------------------------------- 1287

Qy        488 AsnGluGluLysHisProIleIleThrSerSer-----------------LeuLeuGln 501
                ||||||:::|||              ::::::          :::   :::
Db       1288 GAGGAGGAGGAGCACGACGCGGCGGCGGCCGCCGCCGGCGACCGTTGCGGCGTGGCGGAG 1347

Qy        502 AspGlyIleThrThrThrGlnAlaGluIleSerThrSerThrIleSerThrSerProThr 521
                    :::                                  |||:::||||||
Db       1348 CAGGCTCCGAGTAAGCCGGACGACAGTGCTGGCATTGGCATGTCGTCGTCGTCGCCGGCG 1407

Qy        522 AlaGlyAlaSer------------LeuHis------------HisAlaHisAsnPheSer 533
              |||   ||||||            :::|||            ||||||
Db       1408 GCGGCGGCGTCCAGGTCAGTTGGTGTGCACGCCGGAGACCAGCACGCTCAGGCGAGCTTC 1467

Qy        534 PheLeuGlySerPheAsnMetAspAsn--------------------------ThrThr 544
              :::   |||                |||:::                        |||
Db       1468 TACGGCGGCGGCGGCGGCGGCGACGATCCGTTCCAGTGCAGGATCAAGAAGGCGAGGACG 1527

Qy        545 ThrThrValAsp-----------------------HisIleGluAsnAsnAlaLys--- 555
              ||||||   |||                      |||:::         |||
Db       1528 ACGACGGCGGACGAGCCGGCGGCGGCCGCGGCGTTCCACGTCTCCGGCGAGGCCGCCGTG 1587

Qy        556 LysGlnArgAsnAspMetHisLysPheSerProSerSer--------------------- 568
                    |||      :::   ||||||:::   :::
Db       1588 AGCCACCGGGAGCTGCTAATGAAGTTCACTGAGGCCGGCGGCGAGGGCGTGAGAACGGAC 1647

Qy        569 ------IleLeuSerSerValAspMetGluAlaLysAlaArgGluSerSerAsnLysGly 586
                   ::::::      |||   :::
Db       1648 CACCCGCACGTCAACGACGACGACGACGACGTCCCTGGCGGCGCCGGCGGGTACTCGCTG 1707

Qy        587 PheThrAsnProLeuMetAlaAlaTyrAlaMetGlyAspPheGlyArgPheAspProHis 606
              ||||||              ||||||                ||||||        |||
Db       1708 TTCACG------------GCGGCGCAGTACGGGCATCAGTTCGGC------TCGGACCAT 1749

Qy        607 AspGlnGlnMetThrAlaAsnPheHisGlyAsnAsnGlyValSerLeuThrLeuGlyLeu 626
                                       |||      ||||||||||||||||||||||||
```

```
Db       1750 TTCGCGTTCGCCGGGCACGGCGGCGGAGGCGGCGGCGGCGTGTCGCTCACGCTCGGCCTC 1809

Qy        627 ProProSerGluAsnLeuAlaMetProValSerGlnGlnAsnTyrLeuSerAsnAspLeu 646
              |||       :::       |||   |||         :::|||
Db       1810 CCGCACGGCGCCGAC---CAGACGCCGGCGTCG---------TTTCTCATAGGCGCCGGT 1857

Qy        647 GlySerArgSerGluMetGlySerHisTyrAsnArgMetGlyTyrGluAsnIleAspPhe 666
                          |||:::               ||||||:::   ::::::
Db       1858 GCCGGCAGCGACGGCGGCGGCGCGCCGGTGACGACGGCCGGCTACGAC---ATGAACATG 1914

Qy        667 GlnSerGlyAsnLysArgPheProThrGlnLeuLeuProAspPheVal 682
              ||||||     |||           ||||||:::   |||||||||
Db       1915 CAGAGC---ACCAAGTCGTTGGCTGCTCAGCTCATGAGAGACTTCGTG 1959

RESULT 9
BI920940
LOCUS       BI920940                 708 bp    mRNA    linear   EST 10-MAR-2003
DEFINITION  EST540875 potato microtubers, in vitro-grown Solanum tuberosum cDNA
            clone cSTE26N18 5' end, mRNA sequence.
ACCESSION   BI920940
VERSION     BI920940.1  GI:16216968
KEYWORDS    EST.
SOURCE      Solanum tuberosum (potato)
  ORGANISM  Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum.
REFERENCE   1  (bases 1 to 708)
  AUTHORS   van der Hoeven,R., Bezzerides,J., Bachem,C., Visser,R.,
            Karamycheva,S.A., Tsai,J., Van Aken,S., Utterback,T., Chiemingo,A.,
            Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.
  TITLE     Generation of ESTs from in vitro grown microtubers (2001b)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Robin Buell
            The Institute for Genomic Research
            9712 Medical Center Dr, Rockville, MD 20850, USA
            Email: potato-array@tigr.org
            This clone can be obtained from the University of Arizona Genomics
            Institute. Orders can be made through URL:
            http://genome.arizona.edu/orders/
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1. .708
                     /organism="Solanum tuberosum"
                     /mol_type="mRNA"
                     /cultivar="Bintje"
                     /db_xref="taxon:4113"
                     /clone="cSTE26N18"
                     /tissue_type="axillary buds of stem explants; growing
                     sink-tubers"
                     /dev_stage="7, 8 and 10 days"
                     /lab_host="SOLR"
                     /clone_lib="potato microtubers, in vitro-grown"
                     /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
                     XhoI; Tissue supplied by Christian Bachem and Richard
                     Visser (Department of Plant Breeding, Wageningen
                     University, The Netherlands); sequencing by The Institute
                     for Genomic Research. The cSTA libraries will attempt to
                     capture the induction and initiation/initial growth of the
                     tuber in an in vitro system as described in Bachem et al.
                     (Plant Journal, 1996). Small microtubers develop from
                     axillary buds attached to stem explants when placed on a
                     high sucrose medium (10%). Visible morphological changes
                     occur synchronously at day five in the axillary buds. The
                     first library, cSTA (1-20) consists of axillary buds
                     harvested on days 1-3. This targets those genes involved
                     in induction of the microtubers. The following libraries,
                     cSTA (21-40) and cSTA (41-60), capture genes involved in
                     tuber initiation and outgrowth. This library is noted as
                     P3  in Tanksley lab notebooks."
ORIGIN

Alignment Scores:

Pred. No.:              3.01e-88        Length:         708
Score:                  961.00          Matches:        186
Percent Similarity:     94.97%          Conservative:   3
Best Local Similarity:  93.47%          Mismatches:     8
Query Match:            26.97%          Indels:         2
DB:                     4               Gaps:           2

US-10-624-201A-2 (1-688) x BI920940 (1-708)

Qy          1 MetTyrTyrGlnGlyThrSerAspAsnThrAsnIleGlnAlaAspHisGlnGlnArgHis 20
              |||||||||||||||||||||||||||||   |||||||||||||||||||||||||   |||
Db        110 ATGTACTATCAAGGAACCTCGGATAAT---AATATACAAGCTGATCATCAACAACATCAT 166

Qy         21 AsnHisGlyAsnSerAsnAsnAsnAsnIleGlnThrLeuTyrLeuMetAsnProAsnAsn 40
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 AATCATGGGAATAGTAATAATAATAATATTCAGACACTTTATTTGATGAACCCTAACAAT 226

Qy         41 TyrMetGlnGlyTyrThrThrSerAspThr---GlnGlnGlnGlnGlnLeuLeuPheLeu 59
              |||||||||||||||||||||||||||||   |||||||||||||||||||||||||||
Db        227 TATATGCAAGGCTACACTACTTCTGACACACAGCAGCAGCAGCAGTTACTTTTCCTG 286

Qy         60 AsnSerSerProAlaAlaSerAsnAlaLeuCysHisAlaAsnIleGlnHisAlaProLeu 79
              |||||||||||||||   |||||||||||||||||||||||||||||||||||||||||||
Db        287 AATTCTTCACCAGCAGGAAGCAACGCGCTTTGCCATGCGAATATACAACACGCGCCGCTG 346

Qy         80 GlnGlnGlnHisPheValGlyValProLeuProAlaValSerLeuHisAspGlnIleAsn 99
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 CAACAGCAGCACTTTGTCGGTGTGCCTCTTCCGGCAGTAAGTTTGCACGATCAGATCAAT 406

Qy        100 HisHisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnSerGlnGlnValIleVal 119
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 CATCATGGACTTTTACAGCGTATGTGGAACAACCAAGATCAATCTCAGCAGGTGATAGTA 466

Qy        120 ProSerSerThrGlyValSerAlaThrSerCysGlyGlyIleThrThrAspLeuAlaSer 139
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 CCATCGTCGACGGGGGTTTCTGCCACGTCATGTGGCGGGATTACCACGGACTTGGCGTCT 526

Qy        140 GlnLeuAlaPheGlnArgProIleProThrProGlnHisArgGlnGlnGlnGlnGlnGln 159
              |||||||||||||||||||||||||         :::|||||||||||||||||||||||
Db        527 CAATTGGCGTTTCAGAGGCCGATTGTGGTGTCGGAACACCGACAGCAGCAACAACAGCAA 586

Qy        160 GlyGlyLeuSerLeuSerLeuSerProGlnLeuGlnGlnGlnIleSerPheAsnAsnAsn 179
              |||||||||||||||||||||||||||||   |||||||||||||||||||||||||||||
Db        587 GGCGGTCTATCTCTAAGCCTTTCTCCTCAGCAACAACAGCAAATTAGTTTCAATAACAAT 646

Qy        180 IleSerSerSerSerProArgThrAsnAsnValThrIleArgGlyThrLeuAspGly 198
              ||||||:::||||||||||:::   ||||||||||||||||||   |||||||||
Db        647 ATTTCAGCCTCATCACCCAAGGACAATAATGTTACTATTAGAGGAAGATTAGATGGG 703

RESULT 10
AW441343
LOCUS       AW441343                 623 bp    mRNA    linear   EST 18-MAY-2001
DEFINITION  EST310739 tomato fruit red ripe, TAMU Lycopersicon esculentum cDNA
            clone cLEN15B11 5', mRNA sequence.
ACCESSION   AW441343
VERSION     AW441343.1  GI:6976594
KEYWORDS    EST.
SOURCE      Lycopersicon esculentum (tomato)
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1  (bases 1 to 623)
  AUTHORS   Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F.,
            Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S.,
            Ronning,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and
            Giovannoni,J.
  TITLE     Generation of ESTs from tomato fruit tissue
  JOURNAL   Unpublished (1999)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
```

```
          Email: http://www.genome.clemson.edu/orders/index.html
          5 prime sequence.
FEATURES             Location/Qualifiers
     source          1. .623
                     /organism="Lycopersicon esculentum"
                     /mol_type="mRNA"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cLEN15B11"
                     /tissue_type="pericarp"
                     /dev_stage="red ripe (7-20 days post-breaker)"
                     /clone_lib="tomato fruit red ripe, TAMU"
                     /note="Vector: pBlueScript SK(-); Site_1: EcoR1; Site_2:
                     Xho1; supplier: Giovannoni;  Fruit were tagged at the
                     breaker stage (first sign of lycopene accumulation on the
                     blossom end of the fruit) and harvested 7 days
                     post-breaker (fully red-ripe), 10 days post breaker, and
                     20 days post-breaker (over-ripe).  20 day fruit which
                     showed external or internal signs of pathogenesis were
                     discarded.  Fruit were cut in half and the seeds and
                     locules were discarded prior to freezing the pericarp."
ORIGIN

Alignment Scores:
Pred. No.:              3.89e-85        Length:         623
Score:                  930.00          Matches:        177
Percent Similarity:     94.55%          Conservative:   14
Best Local Similarity:  87.62%          Mismatches:     9
Query Match:            26.10%          Indels:         2
DB:                     2               Gaps:           1

US-10-624-201A-2 (1-688) x AW441343 (1-623)

Qy        270 ValGluLeuThrThrAlaGlnArgGlnGluLeuGlnMetLysLysAlaLysLeuLeuAla 289
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 GTTGAACTTACAACAGCTCAAAGACAAGAACTTCAAATGAAGAAAGCAAAGCTTGTTAGC 67

Qy        290 MetLeuGluGluValGluGlnArgTyrArgGlnTyrHisHisGlnMetGlnIleIleVal 309
              ||||||:::|||||||:::|||||||||||||||||||||||||||||||||:::|||
Db         68 ATGCTTGATGAGGTGGATCAAAGGTATAGACAATACCATCACCAAATGCAAATGATTGCA 127

Qy        310 LeuSerPheGluGlnValAlaGlyIleGlySerAlaLysSerTyrThrGlnLeuAlaLeu 329
              ||||||||||||      ||||||||||||:::||||||||||||||||||||||||||
Db        128 ACATCATTTGAGCAAACAACAGGAATTGGATCATCAAAATCATACACACAACTTGCTTTG 187

Qy        330 HisAlaIleSerLysGlnPheArgCysLeuLysAspAlaIleAlaGluGlnValLysAla 349
              |||   ||||||||||||||||||||||||||||||||:::   |||:::|||
Db        188 CACACAATTTCAAAGCAATTTAGATGTTTAAAAGATGCAATTTCTGGGCAAATAAAGGAC 247

Qy        350 ThrSerLysSerLeuGlyGluGluGluGlyLeuGlyGlyLysIleGluGlySerArgLeu 369
              |||:::|||:::|||||||||||||||   :::|||||||||||||||||||:::|||
Db        248 ACAAATAAAACTTTAGGGGAAGAAGAGAACATTGGTGGCAAAATTGAAGGATCAAAGTTG 307

Qy        370 LysPheValAspHisHisLeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGln 389
              ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||
Db        308 AAATTTGTGGATCATCATTTACGCCAACAACGTGCACTACAACAATTAGGGATGATGCAA 367

Qy        390 ProAsnAlaTrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAla 409
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 ACCAATGCATGGAGGCCACAAAGAGGTTTGCCCGAAAGAGCGGTTTCGGTTCTCCGCGCT 427

Qy        410 TrpLeuPheGluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAla 429
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 TGGCTTTTCGAGCATTTTCTTCATCCGTATCCTAAAGATTCAGATAAAATCATGCTTGCT 487

Qy        430 LysGlnThrGlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArg 449
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        488 AAGCAAACAGGGCTAACAAGGAGCCAGGTATCAAATTGGTTCATAAATGCTAGAGTTAGA 547

Qy        450 LeuTrpLysProMetValGluGluMetTyrLeuGluGluValLys------AsnGlnGlu 467
              ||||||||||||||||||||||||||||:::||||||||||||      |||||||||
Db        548 CTATGGAAGCCAATGGTAGAAGAAATGTACATGGAAGAAGTGAAGAAATACAATCAAGAA 607

Qy        468 GlnAsn 469
              ||||||
Db        608 CAAAAT 613

RESULT 11
AI773253
LOCUS       AI773253                553 bp    mRNA    linear   EST 18-MAY-2001
DEFINITION  EST254353 tomato resistant, Cornell Lycopersicon esculentum cDNA
            clone cLER5L8, mRNA sequence.
ACCESSION   AI773253
VERSION     AI773253.1  GI:5271294
KEYWORDS    EST.
SOURCE      Lycopersicon esculentum (tomato)
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1  (bases 1 to 553)
  AUTHORS   D' Ascenzo,M., He,X., Lyman,J., Matern,A.L., Vision,T., Holt,I.E.,
            Liang,F., Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.Y.,
            Bowman,C.L., Nierman,W., Fraser,C.M., Venter,J.C., Tanksley,S.D.,
            Giovannoni,J.J. and Martin,G.B.
  TITLE     Generation of ESTs from Pseudomonas resistant tomato
  JOURNAL   Unpublished (1999)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            5 prime sequence.
FEATURES             Location/Qualifiers
     source          1. .553
                     /organism="Lycopersicon esculentum"
                     /mol_type="mRNA"
                     /cultivar="R11-12 (35S::Pto in Rio Grande x Money Maker)"
                     /db_xref="taxon:4081"
                     /clone="cLER5L8"
                     /tissue_type="leaf"
                     /dev_stage="4-week old"
                     /lab_host="SOLR"
                     /clone_lib="tomato resistant, Cornell"
                     /note="Vector: pBlueScript SK(-); Site_1: EcoR1; Site_2:
                     Xho1; cLER - Tomato Pseudomonas Resistant EST Library.
                     Directionally cloned  cDNAs inserted into pBlueScript
                     SK(-) at 5' end with EcoRI and 3' end  with XhoI site."
ORIGIN

Alignment Scores:
Pred. No.:              3.47e-84        Length:         553
Score:                  920.00          Matches:        175
Percent Similarity:     98.36%          Conservative:   5
Best Local Similarity:  95.63%          Mismatches:     3
Query Match:            25.82%          Indels:         0
DB:                     1               Gaps:           0

US-10-624-201A-2 (1-688) x AI773253 (1-553)

Qy        317 GlyIleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPhe 336
              |||:::|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          3 GGAGTTGGATCAGCAAAATCATACACTCAATTAGCTTTGCATGCAATTTCAAAGCAATTC 62

Qy        337 ArgCysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeuGlyGlu 356
              ||||||||||||||||||:::|||||||||||||||||||||||||||||||||||||
Db         63 AGATGCCTAAAGGATGCAATTAGTGAGCAAGTGAAGGCGACGAGCAAGAGTTTAGGTGAA 122

Qy        357 GluGluGlyLeuGlyGlyLysIleGluGlySerArgLeuLysPheValAspHisHisLeu 376
              :::|||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 GATGAAGGCTTGGGAGGGAAAATTGAAGGCTCAAGACTCAAATTTGTGGACCATCATCTA 182

Qy        377 ArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGlnProAsnAlaTrpArgProGln 396
              ||||||||||||||||||||||||||:::|||||||||||||||||||||||||||||||
Db        183 AGGCAACAACGCGCGCTGCAACAGCTAGGAATGATGCAACCAAATGCTTGGAGACCCCAA 242
```

```
Qy        397 ArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeuPheGluHisPheLeu 416
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 AGAGGTTTACCTGAAAGAGCTGTCTCTGTCCTTCGTGCTTGGCTTTTCGAGCATTTTCTT 302

Qy        417 HisProTyrProLysAspSerAspLysIleMetLeuAlaLysGlnThrGlyLeuThrArg 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 CATCCTTACCCCAAAGACTCAGACAAAATCATGCTTGCTAAGCAAACGGGGCTAACAAGG 362

Qy        437 SerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLysProMetValGlu 456
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 AGCCAGGTCTCTAACTGGTTCATTAATGCTCGAGTTCGATTATGGAAGCCAATGGTTGAA 422

Qy        457 GluMetTyrLeuGluGluValLysAsnGlnGluGlnAsnSerThrAsnThrSerGlyAsp 476
              ||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||||
Db        423 GAAATGTACTTGGAAGAAGTGAAGAATCAAGAACAAAACAGTAGTAATACTTCAGGAGAT 482

Qy        477 AsnLysAsnLysGluThrAsnIleSerAlaProAsnGluGluLysHisProIleIleThr 496
              ||||||||||||||||||      ||||||||||||||||||||   ||||||||||||
Db        483 AACAAAAACAAAGAGACGAAATTTAGTGCTCCAAATGAAGAGAAACAACCCATTATTACT 542

Qy        497 SerSerLeu 499
              |||||||||
Db        543 AGCAGCTTA 551

RESULT 12
CV241175/c
LOCUS       CV241175                827 bp    mRNA    linear   EST 22-SEP-2004
DEFINITION  WS02511.B21_G23 PT-MB-N-A-15 Populus balsamifera subsp. trichocarpa
            cDNA clone WS02511_G23 3', mRNA sequence.
ACCESSION   CV241175
VERSION     CV241175.1  GI:52494150
KEYWORDS    EST.
SOURCE      Populus balsamifera subsp. trichocarpa (Populus trichocarpa)
  ORGANISM  Populus balsamifera subsp. trichocarpa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Malpighiales; Salicaceae; Saliceae; Populus.
REFERENCE   1  (bases 1 to 827)
  AUTHORS   Ralph,S., Cooper,D., Kolosova,N., Oddy,C., Butterfield,Y.,
            Kirkpatrick,R., Liu,J., Palmquist,D., Stott,J., Barber,S., Yang,G.,
            Babakaiff,R., Brown-John,M., Chand,S., Featherstone,R., Masson,A.,
            Mayo,M., Moran,J., Olson,T., Wong,D., Ritland,C.E., Siddiqui,A.,
            Holt,R., Jones,S., Marra,M., Ellis,B.E., Douglas,C., Ritland,K. and
            Bohlmann,J.
  TITLE     The poplar transcriptome: Analysis of expressed sequence tags from
            multiple cDNA libraries
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Joerg Bohlmann
            Genome BC forest genomics program
            University of British Columbia
            UBC Biotechnology Laboratory, 6174 University Boulevard, Rm. 237,
            Vancouver, British Columbia, Canada, V6T 1Z3
            Tel: 1-604-822-0282
            Fax: 1-604-822-6097
            Email: bohlmann@interchange.ubc.ca
            Plate: WS02511  row: G  column: 23
            High quality sequence stop: 827.
FEATURES             Location/Qualifiers
     source          1. .827
                     /organism="Populus balsamifera subsp. trichocarpa"
                     /mol_type="mRNA"
                     /cultivar="Wild clone"
                     /sub_species="trichocarpa"
                     /db_xref="taxon:3694"
                     /clone="WS02511_G23"
                     /sex="Male"
                     /lab_host="E. coli DH10B T1 phage resistant cells"
                     /clone_lib="PT-MB-N-A-15"
                     /note="Vector: pBluescript II SK (+) XR; Site_1: EcoRI (5'
                     end of cDNA); Site_2: XhoI (3' end of cDNA); Terminal
                     vegetative buds from 20 year old trees harvested near
                     Corvallis, Oregon on September 19th, 2001. cDNA was
                     prepared from 5 micrograms of mRNA and directionally
                     ligated into the pBluescript II SK (+) XR vector using the
                     pBluescript II XR cDNA Library Construction Kit according
                     to manufacturer's instructions with modifications
                     (Stratagene). Plasmid DNA was then transformed by
                     electroporation into DH10B cells (Invitrogen) for
                     propagation. Normalization was applied according to
                     published methods [Bonaldo M.F. et al. (1996) Genome
                     Research 6(9):791] in order to reduce the abundance of
                     highly expressed transcripts."
ORIGIN

Alignment Scores:
Pred. No.:                1.22e-81        Length:          827
Score:                    898.00          Matches:         184
Percent Similarity:       73.72%          Conservative:    32
Best Local Similarity:    62.80%          Mismatches:      55
Query Match:              25.20%          Indels:          22
DB:                       7               Gaps:            4

US-10-624-201A-2 (1-688) x CV241175 (1-827)

Qy        155 GlnGlnGlnGlnGlnGlyGlyLeuSerLeuSerLeuSerProGlnLeuGlnGlnGlnIle 174
              |||  |||        |||      :::  :::|||
Db        820 CAGGCTCAAACTGTATCTGGTGAAGATATAAGGGTGTCC---------------------- 782

Qy        175 SerPheAsnAsnAsnIleSerSerSerSerProArgThrAsnAsnValThrIleArgGly 194
                      ::::::||||||   |||      ||||||   |||
Db        781 ------GGTGGGTCGGTGTCATCAGGCTCGGGTGTGACAAATGGGGTTCTG--------- 737

Qy        195 ThrLeuAspGlySerSerSerAsnMetValLeuGlySerLysTyrLeuLysAlaAlaGln 214
                                  ::::::|||  ||||||||||||||||||||||
Db        736 -----------GGTATGCAAGGGGTGTTGCTGAGCTCAAAGTACTTGAAGGCTGCTCAA 689

Qy        215 GluLeuLeuAspGluValValAsnIleValGlyLysSerIleLysGlyAspAspGlnLys 234
              ||||||||||||||||||||::::::          ||||||  :::      |||
Db        688 GAGCTACTTGATGAGGTTGTTAGCGTGAATAATAATGATATTAAGAGTGAATTGTCAAAG 629

Qy        235 LysAspAsnSerMet------AsnLysGluSerMetProLeuAlaSerAspValAsnThr 252
              :::  |||  :::      |||      :::      :::         :::
Db        628 AGGAGTAATGGGATTGGTAGTAACACTAGTAACAAAGTGGTTGGAGAATCATTAGCCGGA 569

Qy        253 AsnSerSerGlyGlyGlyGluSerSerSerArgGlnLysAsnGluValAlaValGluLeu 272
                    ||||||||||||||  |||  ::::::                  ||||||
Db        568 GAAGGGTCTGGTGGAGGAGAAGTAAGTGGGAAGCGT-----------GGACCGGAGCTT 521

Qy        273 ThrThrAlaGlnArgGlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeuGlu 292
              :::||||||:::|||||||||:::||||||||||||||||||:::::|||||:::
Db        520 TCCACTGCCGAGAGGCAGGAAATTCAGATGAAGAAGGCTAAACTTATCAGCATGCTTGAT 461

Qy        293 GluValGluGlnArgTyrArgGlnTyrHisHisGlnMetGlnIleIleValLeuSerPhe 312
              ||||||||||||||||||||||||||||||||||||||||||:::::   ||||||
Db        460 GAGGTGGAGCAAAGGTACAGGCAGTATCATCACCAGATGCAGATTGTGATTTCCTCGTTT 401

Qy        313 GluGlnValAlaGlyIleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaIle 332
              ||||||      ||||||||||||||||:::||||||   ||||||||      |||
Db        400 GAGCAAGCAGTAGGAATTGGTTCCGCTAAGACATATACAGCCCTTGCGCTGAAAACAATC 341

Qy        333 SerLysGlnPheArgCysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLys 352
              ||||||||||||||||||||||||||||||||||||||      |||:::||||||   :::|||
Db        340 TCCAAGCAGTTTAGGTGCTTGAAGGATGCAATAACAGGTCAAATTAAAGCTGCAAACAAA 281

Qy        353 SerLeuGlyGluGluGluGlyLeuGlyGlyLysIleGluGlySerArgLeuLysPheVal 372
              ||||||||||||||:::   ||||||||||||||||||||||||||||||||||||||||
Db        280 AGCTTAGGTGAAGAGGATTGCTTGGGAGGAAAGATTGAAGGTTCAAGGCTCAAATTTGTC 221

Qy        373 AspHisHisLeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGlnProAsnAla 392
              ||||||||||||||||||||||||||||||||||:::||||||:::|||    ||||||
Db        220 GATCATCACCTTAGGCAACAGCGTGCACTTCAGCAGTTGGGAATGATCCAGCACAATGCT 161

Qy        393 TrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeuPhe 412
              ||||||||||||||||||||||||||||||||||:::|||||||:::||||||||||||||||||
Db        160 TGGAGACCACAGAGAGGATTGCCTGAAAGATCAGTTTCACTTCTCCGTGCTTGGCTCTTC 101
```

```
Qy        413 GluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGlnThr 432
              ||||||||||||||||||||||||||||||||||||||||   ||||||||||||||||||
Db        100 GAACACTTTCTCCACCCCTATCCCAAGGATTCAGACAAGCACATGCTCGCAAAACAAACA 41

Qy        433 GlyLeuThrArgSerGlnValSerAsnTrpPheIleAsn 445
              |||||||||||||||||||||||||||||||||||||
Db         40 GGACTCACGAGGAGCCAGGTGTCTAATTGGTTCATAAAC 2

RESULT 13
BI921008
LOCUS       BI921008                789 bp    mRNA    linear   EST 10-MAR-2003
DEFINITION  EST540943 potato microtubers, in vitro-grown Solanum tuberosum cDNA
            clone cSTE27I21 5' end, mRNA sequence.
ACCESSION   BI921008
VERSION     BI921008.1  GI:16217036
KEYWORDS    EST.
SOURCE      Solanum tuberosum (potato)
  ORGANISM  Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum.
REFERENCE   1  (bases 1 to 789)
  AUTHORS   van der Hoeven,R., Bezzerides,J., Bachem,C., Visser,R.,
            Karamycheva,S.A., Tsai,J., Van Aken,S., Utterback,T., Chiemingo,A.,
            Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.
  TITLE     Generation of ESTs from in vitro grown microtubers (2001b)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Robin Buell
            The Institute for Genomic Research
            9712 Medical Center Dr, Rockville, MD 20850, USA
            Email: potato-array@tigr.org
            This clone can be obtained from the University of Arizona Genomics
            Institute. Orders can be made through URL:
            http://genome.arizona.edu/orders/
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1. .789
                     /organism="Solanum tuberosum"
                     /mol_type="mRNA"
                     /cultivar="Bintje"
                     /db_xref="taxon:4113"
                     /clone="cSTE27I21"
                     /tissue_type="axillary buds of stem explants; growing
                     sink-tubers"
                     /dev_stage="7, 8 and 10 days"
                     /lab_host="SOLR"
                     /clone_lib="potato microtubers, in vitro-grown"
                     /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
                     XhoI; Tissue supplied by Christian Bachem and Richard
                     Visser (Department of Plant Breeding, Wageningen
                     University, The Netherlands); sequencing by The Institute
                     for Genomic Research. The cSTA libraries will attempt to
                     capture the induction and initiation/initial growth of the
                     tuber in an in vitro system as described in Bachem et al.
                     (Plant Journal, 1996). Small microtubers develop from
                     axillary buds attached to stem explants when placed on a
                     high sucrose medium (10%). Visible morphological changes
                     occur synchronously at day five in the axillary buds. The
                     first library, cSTA (1-20) consists of axillary buds
                     harvested on days 1-3. This targets those genes involved
                     in induction of the microtubers. The following libraries,
                     cSTA (21-40) and cSTA (41-60), capture genes involved in
                     tuber initiation and outgrowth. This library is noted as
                     P3  in Tanksley lab notebooks."
ORIGIN

Alignment Scores:
Pred. No.:              2.31e-81        Length:         789
Score:                  895.00          Matches:        168
Percent Similarity:     99.41%          Conservative:   0
Best Local Similarity:  99.41%          Mismatches:     1
Query Match:            25.12%          Indels:         0
DB:                     4               Gaps:           0

US-10-624-201A-2 (1-688) x BI921008 (1-789)

Qy          1 MetTyrTyrGlnGlyThrSerAspAsnThrAsnIleGlnAlaAspHisGlnGlnArgHis 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        282 ATGTACTATCAAGGAACCTCGGATAATACTAATATACAAGCTGATCATCAACAACATCAT 341

Qy         21 AsnHisGlyAsnSerAsnAsnAsnAsnIleGlnThrLeuTyrLeuMetAsnProAsnAsn 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        342 AATCATGGGAATAGTAATAATAATAATATTCAGACACTTTATTTGATGAACCCTAACAAT 401

Qy         41 TyrMetGlnGlyTyrThrThrSerAspThrGlnGlnGlnGlnGlnLeuLeuPheLeuAsn 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        402 TATATGCAAGGCTACACTACTTCTGACACACAGCAGCAGCAGCAGTTACTTTTCCTGAAT 461

Qy         61 SerSerProAlaAlaSerAsnAlaLeuCysHisAlaAsnIleGlnHisAlaProLeuGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        462 TCTTCACCAGCAGCAAGCAACGCGCTTTGCCATGCGAATATACAACACGCGCCGCTGCAA 521

Qy         81 GlnGlnHisPheValGlyValProLeuProAlaValSerLeuHisAspGlnIleAsnHis 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        522 CAGCAGCACTTTGTCGGTGTGCCTCTTCCGGCAGTAAGTTTGCACGATCAGATCAATCAT 581

Qy        101 HisGlyLeuLeuGlnArgMetTrpAsnAsnGlnAspGlnSerGlnGlnValIleValPro 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        582 CATGGACTTTTACAGCGCATGTGGAACAACCAAGATCAATCTCAGCAGGTGATAGTACCA 641

Qy        121 SerSerThrGlyValSerAlaThrSerCysGlyGlyIleThrThrAspLeuAlaSerGln 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        642 TCGTCGACGGGGGTTTCTGCCACGTCATGTGGCGGGATCACCACGGACTTGGCGTCTCAA 701

Qy        141 LeuAlaPheGlnArgProIleProThrProGlnHisArgGlnGlnGlnGlnGlnGlnGly 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        702 TTGGCGTTTCAGAGGCCGATTCCGACACCACAACACCGACAGCAGCAACAACAGCAAGGC 761

Qy        161 GlyLeuSerLeuSerLeuSerProGln 169
              |||||||||||||||||||||||||||
Db        762 GGTCTATCTCTAAGCCTTTCTCCTCAG 788

RESULT 14
CB292855
LOCUS       CB292855                850 bp    mRNA    linear   EST 28-FEB-2003
DEFINITION  UCRCS01_05af07_b1 Washington Navel orange cold acclimated flavedo &
            albedo cDNA library Citrus sinensis cDNA clone UCRCS01_05af07, mRNA
            sequence.
ACCESSION   CB292855
VERSION     CB292855.1  GI:28618312
KEYWORDS    EST.
SOURCE      Citrus sinensis
  ORGANISM  Citrus sinensis
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Sapindales; Rutaceae; Citrus.
REFERENCE   1  (bases 1 to 850)
  AUTHORS   Close,T.J., Collura,K., Fenton,R.D., Feuerbacher,O., Kim,H.R.,
            Kudrna,D., Wanamaker,S., Wing,R. and Yu,Y.
  TITLE     Development of EST Resources and New Genetic Markers for California
            Citrus
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Timothy Close
            Department of Botany & Plant Sciences, University of California
            Riverside, CA, 92521-0124
            Tel: 9097873318
            Fax: 9097874437
            Email: timothy.close@ucr.edu
            Seq primer: T7.
FEATURES             Location/Qualifiers
     source          1. .850
                     /organism="Citrus sinensis"
                     /mol_type="mRNA"
                     /cultivar="Washington navel"
                     /db_xref="taxon:2711"
                     /clone="UCRCS01_05af07"
```

/tissue_type="Rind containing flavedo and albedo"
/dev_stage="Mature fruit"
/lab_host="E. coli TJC121"
/clone_lib="Washington Navel orange cold acclimated flavedo & albedo cDNA library"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid; Site_1: EcoRI; Site_2: XhoI; Plants were grown in the field at University of California, Riverside Agricultural Operations since 1983. The scion was Washington Navel orange and the rootstock Carizzo Citrange. Tissue from mature fruit was harvested at mid-day in January 2002 during a cold spell, when pre-dawn temperatures were approximately -2 to 2 degree C. Approximately 2 cm median sections of the rind were excised in the field from several fruits, then wrapped in aluminum foil and frozen quickly in dry ice. Total RNA was extracted using a phenol extraction procedure described in J. Japanaese Soc. Hort. Sci. (1996) 64:809-814. Poly(A) RNA was purified, a cDNA library was made, and 1 million primary lambda cDNA clones were in vivo excised to give a population of pBluescript SK(-) phagemids. All steps to this point were performed in the TJ Close lab at the University of California, Riverside (Fenton). Phagemids were plated, plasmid DNA purified, cDNA clones archived, and DNA sequences determined bi-directionally using an ABI3530 at the Arizona Genomics Institute, University of Arizona (Collura, Feuerbacher, Kim, Kudrna, Wing, Yu). Chromatogram files were transmitted to UC Riverside (by Yu), then processed at UC Riverside (by Wanamaker) using the HarvEST pipeline (http://harvest.ucr.edu) to remove vector and cloning oligo sequences and various contaminants, and to trim to a high quality region. Sequences that retained a phred 17 region of at least 100 bases were deposited to GenBank."

ORIGIN

Alignment Scores:

Pred. No.:	4.71e-81	Length:	850
Score:	892.50	Matches:	183
Percent Similarity:	73.59%	Conservative:	26
Best Local Similarity:	64.44%	Mismatches:	50
Query Match:	25.05%	Indels:	25
DB:	6	Gaps:	3

US-10-624-201A-2.(1-688) x CB292855 (1-850)

```
Qy        272 LeuThrThrAlaGlnArgGlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeu 291
              |||:::||||||:::|||||||||:::|||||||||||||||||              ||||||
Db         17 CTCTCCACGGCAGAGAGGCAAGAAATTCAGATGAAGAAAGCTCATTAAAGTAATATGCTT 76

Qy        292 GluGluValGluGlnArgTyrArgGlnTyrHisHisGlnMetGlnIleIleValLeuSer 311
              :::|||||||||||||||||        |||||||||||||||||::::::   |||
Db         77 GATGAGGTGGAGCAAAGGTACCCGTTGAACCATCACCAGATGCAGATAGTGATTTCATCA 136

Qy        312 PheGluGlnValAlaGlyIleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAla 331
              |||||||||   |||||||||   |||||||||:::||||||   |||||||||
Db        137 TTTGAACAAGCAGCAGGAATTGAGTCAGCAAAGACATACACAGCACTTGCTCTCAAGACA 196

Qy        332 IleSerLysGlnPheArgCysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSer 351
              |||||||||||||||||||||||||||||||||      |||:::||||||   :::
Db        197 ATTTCAAAGCAATTCCGGTGCTTGAAAGATGCAATTACAGGCCAAATTAAAGCGGCAAAT 256

Qy        352 LysSerLeuGlyGluGluGluGlyLeuGlyGlyLysIleGluGlySerArgLeuLysPhe 371
              |||   ||||||||||||:::   :::|||   ||||||||||||||||||||||||||
Db        257 AAGATGCTGGGTGAGGAGGATTGCGTCGGGAGCAAGATTGAAGGCTCCAGGCTCAAATTT 316

Qy        372 ValAspHisHisLeuArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGlnProAsn 391
              |||||||||||||||||||||||||||||||||:::||||||:::|||   |||
Db        317 GTCGACCACCATCTTCGTCAACAAAGGGCTCTTCAACAATTGGGAATGATCCAGCACAAT 376

Qy        392 AlaTrpArgProGlnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeu 411
              |||||||||||||||||||||||||||||||:::|||||||||||||||||||||||||
Db        377 GCTTGGAGACCCCAAAGAGGATTGCCTGAAAGATCTGTCTCAGTTCTTCGCGCTTGGCTC 436

Qy        412 PheGluHisPheLeuHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGln 431
              |||||||||||||||||||||||||||||||||||||||||||||   |||||||||||||
Db        437 TTCGAACACTTCCTTCACCCGTATCCGAAAGATTCGGACAAGCAGATGCTGGCAAAACAA 496

Qy        432 ThrGlyLeuThrArgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrp 451
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 ACAGGGCTTACCAGGAGCCAGGTGTCTAATTGGTTCATAAATGCACGAGTTCGGCTGTGG 556

Qy        452 LysProMetValGluGluMetTyrLeuGluGluValLysAsnGlnGluGlnAsn------ 469
              |||||||||||||||||||||||||||||||||:::|||:::|||||||||||||
Db        557 AAACCAATGGTTGAGGAAATGTACTTGGAGGAAATTAAAGATCAAGAACAGAACGGATCA 616

Qy        470 ----------------------------------------SerThrAsnThrSerGly 475
                                                       |||   :::|||:::
Db        617 GAGGACAAAACAAGCAAGAGTGAGCACAATGAGGATTCGGCATCTAAGTCCACTGCTGCA 676

Qy        476 AspAsnLysAsn------LysGluThrAsnIleSerAlaProAsnGluGluLysHisPro 493
                  ||||||      |||||||||      |||      :::
Db        677 CAAGAGAAAAATCTCGTGAAGGAAACTCAAAACTCGAAGAGTTTCAAATCCAGTGAGGAT 736

Qy        494 IleIleThrSerSerLeuLeuGlnAspGlyIleThrThrThrGlnAlaGluIleSerThr 513
                 :::|||:::                                   :::   ||||||
Db        737 AATTTGACAAACCAG--------------------AATGTTCATTCTATGATTTCAATG 775

Qy        514 SerThrIleSerThrSerProThrAlaGlyAlaSerLeuHisHisAlaHisAsnPheSer 533
              |||||||||:::|||||||||      |||   :::      :::         ||||||
Db        776 TCCACAATTGCAACATCTCCTATCGGAGGAGGGAATGCGAGAAACCATTCAGGATTTTCA 835

Qy        534 PheLeuGlySer 537
                 :::||||||
Db        836 CTCATTGGATCA 847
```

RESULT 15
BM535737
LOCUS BM535737 752 bp mRNA linear EST 20-FEB-2002
DEFINITION EST588759 tomato breaker fruit Lycopersicon esculentum cDNA clone cLEG67P6 5' end, mRNA sequence.
ACCESSION BM535737
VERSION BM535737.1 GI:18814614
KEYWORDS EST.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 752)
AUTHORS Alcala,J., Vrebalov,J., White,R., Vision,T., Karamycheva,S.A., Tsai,J., Bougri,O., Kirkness,E., Utterback,T., Van Aken,S., Ronning,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
TITLE Generation of ESTs from tomato fruit tissue, breaker stage (2002)
JOURNAL Unpublished (2002)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
This clone is available through the Clemson University Genomics Institute
Seq primer: T3.
FEATURES Location/Qualifiers
source 1. .752
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEG67P6"
/tissue_type="Pericarp"
/dev_stage="breaker"
/lab_host="SOLR"
/clone_lib="tomato breaker fruit"
/note="Vector: pBluescriptSKmCUadapt; Site_1: EcoR1;

Site_2: XhoI; supplier: Boyce Thompson Institute; sequencing: The Institute for Genomic Research. Fruit were harvested at the breaker stage (first sign of lycopene accumulation on the blossom end of fruit). Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."

ORIGIN

Alignment Scores:

Pred. No.:	7.91e-81	Length:	752
Score:	889.50	Matches:	188
Percent Similarity:	82.47%	Conservative:	19
Best Local Similarity:	74.90%	Mismatches:	27
Query Match:	24.96%	Indels:	18
DB:	4	Gaps:	4

US-10-624-201A-2 (1-688) x BM535737 (1-752)

```
Qy        237 AsnSerMetAsnLysGluSerMetProLeuAlaSerAspValAsnThrAsnSerSerGly 256
              :::   :::      |||   :::   ||||||||||||
Db          7 CACAAATTGGAGAATGAATTAATATCTTTGGCTAGTGAT---------------------- 45

Qy        257 GlyGlyGluSerSerSerArgGlnLysAsnGluValAlaValGluLeuThrThrAlaGln 276
                     |||:::   |||||||||       ||||||||||||||||||||
Db         46 ---GATATTGAAAGTAACACCCAAAAAAATAGT---GGTGTTGAACTTACAACAGCTCAA 99

Qy        277 ArgGlnGluLeuGlnMetLysLysAlaLysLeuLeuAlaMetLeuGluGluValGluGln 296
              ||||||||||||||||||||||||||||||:::::|||||:::|||||:::|||
Db        100 AGACAAGAACTTCAAATGAAGAAAGCAAAGCTTGTTAGCATGCTTGATGAGGTGGATCAA 159

Qy        297 ArgTyrArgGlnTyrHisHisGlnMetGlnIleIleValLeuSerPheGluGlnValAla 316
              |||||||||||||||||||||||||:::|||      ||||||||||||
Db        160 AGGTATAGACAATACCATCACCAAATGCAAATGATTGCAACATCATTTGAGCAAACAACA 219

Qy        317 GlyIleGlySerAlaLysSerTyrThrGlnLeuAlaLeuHisAlaIleSerLysGlnPhe 336
              ||||||||||||:::|||||||||||||||||||||||||||   ||||||||||||||
Db        220 GGAATTGGATCATCAAAATCATACACACAACTTGCTTTGCACACAATTTCAAAGCAATTT 279

Qy        337 ArgCysLeuLysAspAlaIleAlaGluGlnValLysAlaThrSerLysSerLeu-GlyGl 356
              ||||||||||||||||||:::   |||:::|||   |||:::|||:::||| |||||
Db        280 AGATGTTTAAAAGATGCAATTTCTGGGCAAATAAAGGACACAAATAAAACTTTAGGGGGA 339

Qy        356 uGluGluGlyLeuGlyGlyLysIleGluGlySerArgLeu-LysPheValAspHisHisL 376
              |||||||   :::||||||:::|||||||||||||:::||| ||||||||||||||||||
Db        340 AGAAGAGAACATTGGTGGGCAAATTGAAGGATCAAAGTTGAAAATTTGTGGATCATCATT 399

Qy        376 euArgGlnGlnArgAlaLeuGlnGlnIleGlyMetMetGlnProAsnAlaTrpArgProG 396
              |||||||||||||||||||||||:::|||||||||||   ||||||||||||||||
Db        400 TACGCCAACAACGTGCACTACAACAATTAGGGATGATGCAAACCAATGCATGGAGGCCAC 459

Qy        396 lnArgGlyLeuProGluArgAlaValSerValLeuArgAlaTrpLeuPheGluHisPheL 416
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        460 AAAGAGGTTTGCCCGAAAGAGCGGTTTCGGTTCTCCGCGCTTGGCTTTTCGAGCATTTTC 519

Qy        416 euHisProTyrProLysAspSerAspLysIleMetLeuAlaLysGlnThrGlyLeuThrA 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        520 TTCATCCGTATCCTAAAGATTCAGATAAAATCATGCTTGCTAAGCAAACAGGGCTAACAA 579

Qy        436 rgSerGlnValSerAsnTrpPheIleAsnAlaArgValArgLeuTrpLysProMetValG 456
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        580 GGAGCCAGGTATCAAATTGGTTCATAAATGCTAGAGTTAGACTATGGAAGCCAATGGTAG 639

Qy        456 luGluMetTyrLeuGluGluValLys------AsnGlnGluGlnAsnSerThrAsnThrS 474
              ||   ||||||:::|||||||||||||      ||||||||||||||||
Db        640 AAGA-ATGTACATGGAAGAAGTGAAGAAAAACAATCAAGAACAAAATGGTTTA------- 691

Qy        474 erGlyAspAsnLysAsnLysGluThrAsn 483
                 |||   |||      |||   |||
Db        692 -----GATCAAAAAATTATTGAACCTAAT 715
```

Search completed: July 29, 2005, 18:54:24
Job time : 5786 secs

This Page Blank (uspto)